Quicker/Cheaper Stand Assessments[1]

by

Thomas E. Burk[2]

June 2004
Originally Published October 1999

Staff Paper Series No. 139
Revision 2

Department of Forest Resources

College of Food, Agricultural and Natural Resource Sciences
University of Minnesota
St. Paul, Minnesota

[1]Research supported by Minnesota Agricultural Experiment Station under project MN 42-044. Published as paper no. 994420139 of the Minnesota Agricultural Experiment Station.

[2]The author is Professor, Department of Forest Resources, College of Food, Agricultural and Natural Resource Sciences, University of Minnesota, St. Paul, MN.

For more information about the Department of Forest Resources and its teaching, research, and outreach programs, contact the department at:

Department of Forest Resources
University of Minnesota
115 Green Hall
1530 Cleveland Avenue North
St. Paul, MN 55108-6112
Ph: 612.624.3400
Fax: 612.625.5212
Email: frweb@umn.edu
http://www.forestry.umn.edu/publications/staffpapers/index.html

Table of Contents

Abstract

Sampling and measurement techniques are available for reducing the effort/cost associated with standard timber stand assessments depending on the quantities of interest as well as the volume table(s) being used. This paper will introduce and consider the application of standard variable-radius plot sampling, double (two-phase) sampling, a new technique called Big-BAF sampling, as well as assorted shortcut measurement and observation techniques as they apply to Lake States forestry practice. The relative efficiencies of the techniques from both cost and precision viewpoints are also discussed. The paper is written for individuals who supervise or plan or conduct timber stand assessments for their organization. A basic understanding of the conduct of timber stand assessments using standard methods (e.g. line-plot cruises) is assumed. The paper could be the basis for a workshop; hands-on activities, including a field exercise, are suggested. Spreadsheet code implementing many of the calculations presented in the paper is also available at <http://mallit.fr.umn.edu/burk/quick_cheap/>.

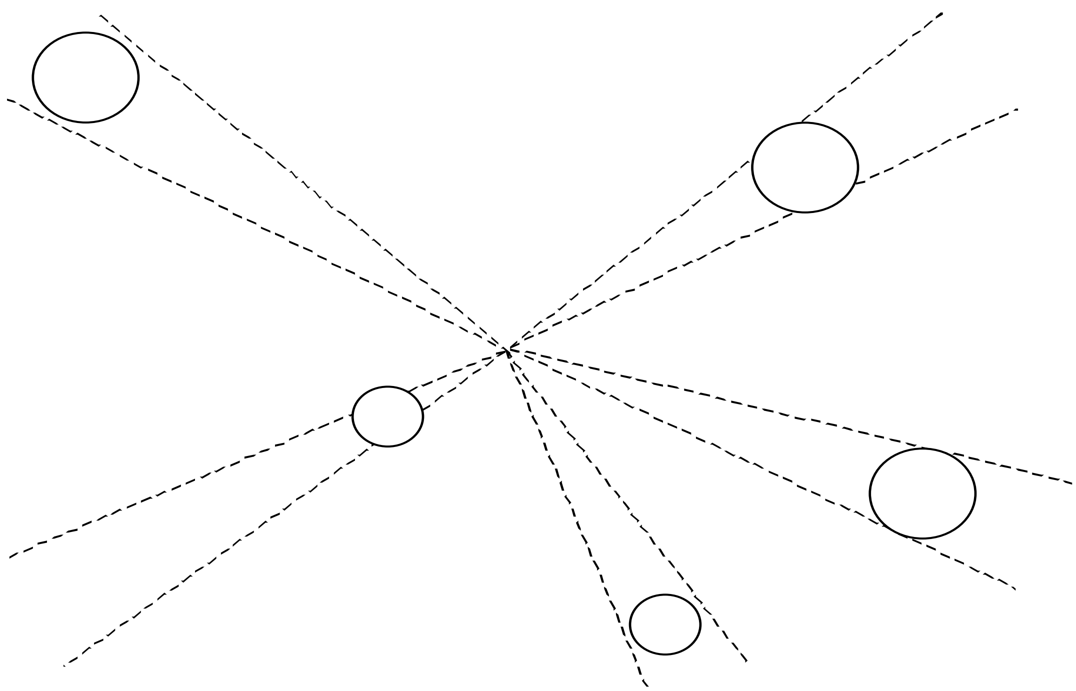
Sampling for timber attributes with variable-radius plots – A review

Why variable-radius plots?

Theory proves that sampling error will be minimized by selecting elements for inclusion in the sample with probability proportional to the variable of interest. For the variable-radius plot type we'll consider, trees at a sample location are selected in proportion to tree basal area. Basal area is typically closely related to volume (at least more so than is tree number), the primary variable of interest in timber stand assessments.

Tree tally at a sample location

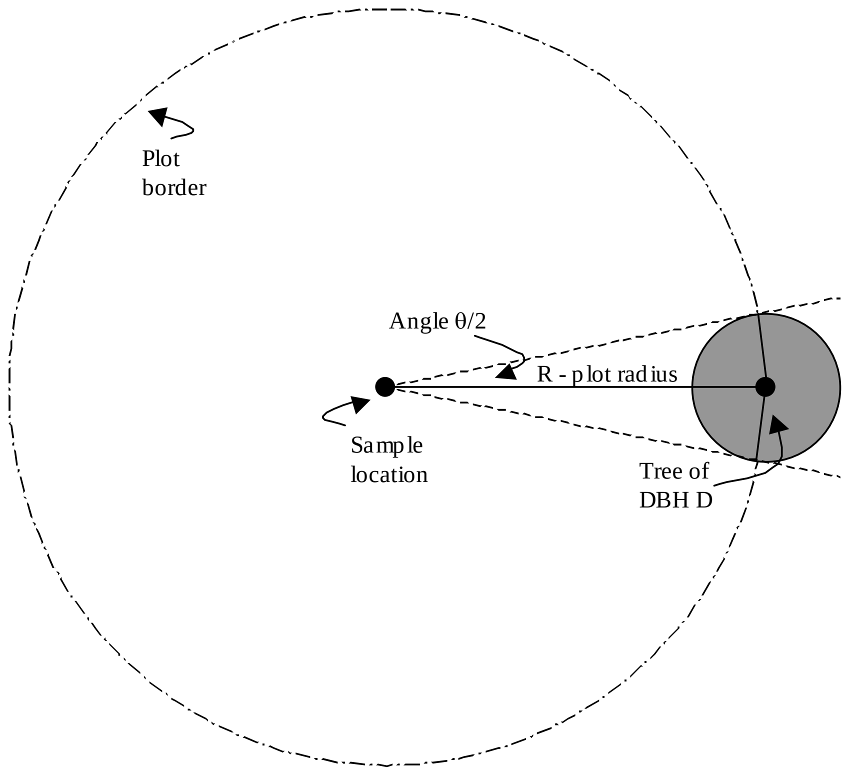
An angle gauge, projecting a fixed, horizontal angle, is used to determine which trees are "in" at a sample location. The tally rule is illustrated below; the schematic is a view from overhead with tree cross-sections at breast height shown. The dashed lines represent the fixed, horizontal angle from the sample location.



A prism works similarly by refracting light rays through an angle equal to that of the fixed horizontal angle; "in" trees are those where the true and false (as seen through the prism) images of the tree at breast height overlap.

Relationship between plot size (for a tree) and tree size

The schematic below shows a tree cross-section at breast height for a tree that is at a distance from a sample location such that it is on the border of its plot; again the dashed lines represent the fixed, horizontal angle. The distance R (feet) is thus the radius of the plot for a tree of DBH D (inches).



Taking q = 2 sin (θ/2), a constant for a given horizontal angle, and watching units gives

$$R = \frac{D}{12q}$$

which means that plot size <u>in acres</u> for any tree i is

$$\begin{aligned} a_i &= \frac{\pi \cdot D_i^2}{43560 \cdot 144 \cdot q^2} \\ &= \frac{\mathrm{B_i}}{10890 \cdot \mathrm{q}^2} \\ &= \frac{\mathrm{B_i}}{BAF} \end{aligned} \qquad [1\text{-}1]$$

where B is tree basal area (square feet – 0.005454DBH^2) and BAF is “basal area factor.” We will use BAF rather than θ to identify an instrument’s constant angle. BAF is the number of square feet per acre that each “in” tree represents at a sample location.

Per acre timber characteristics

In a timber stand assessment we are normally interested in stand attributes like volume per acre, basal area per acre, and stems per acre. To get per acre estimates we need to expand tree characteristics according to a tree's plot size. If there are t trees on a plot ("in" at a sample location) and v represents the volume of a tree (from a table or equation), then volume per acre would be

$$\sum_{i=1}^{t} \frac{v_i}{a_i} = BAF \sum_{i=1}^{t} \frac{v_i}{B_i}$$

The volume to basal area ratio, VBAR, plays a key role in variable-radius plot sampling. If there are n sample locations total with t_j "in" trees at the jth sample location then

$$\text{volume per acre} = \frac{BAF}{n} \sum_{j=1}^{n} \sum_{i=1}^{t_j} \frac{v_i}{B_i} \qquad [1\text{-}2]$$

and likewise

$$\text{stems per acre} = \frac{BAF}{n} \sum_{j=1}^{n} \sum_{i=1}^{t_j} \frac{1}{B_i} \qquad [1\text{-}3]$$

$$\text{basal area per acre} = \frac{BAF}{n} \sum_{j=1}^{n} t_j = \frac{BAF}{n} T \qquad [1\text{-}4]$$

where T is the total number of "in" trees at all n sample locations.

If interest lies in an estimate for some class of tree, the same formulae can be used with the second sum only taken across trees in the class of interest. To find the stock table entry for 8-inch DBH aspen one would sum the volume to basal area ratios for any such trees tallied, multiply by BAF and divide by n. To find the stand table entry for the same trees one would multiply 2.865 (inverse basal area for an 8-inch DBH tree) by the number of such trees tallied and again multiply by BAF and divide by n.

Computing a sampling error

From a statistical perspective, it is the variability between observations at sample locations that is relevant. To compute a sampling error requires separate tally by sample location and computation of the variable(s) of interest sample unit by sample unit. For n sample units taken at random where V_j is the volume per acre observation at location j, the variance of the observations is estimated by

$$s_V^2 = \frac{\sum_{j=1}^{n} V_j^2 - \frac{\left(\sum_{j=1}^{n} V_j\right)^2}{n}}{n-1} \qquad [1\text{-}5]$$

The coefficient of variation of the observations is estimated by

$$CV_V = 100 \cdot \frac{s_V}{\sum_{j=1}^{n} V_j / n} \qquad [1\text{-}6]$$

The sampling error, or standard error of the estimate, if the estimator is the sample average (as suggested above), is estimated by

$$s_{\bar{V}} = \sqrt{s_V^2 / n} \qquad [1\text{-}7]$$

What BAF to use?

Choosing a BAF is like choosing a plot size for fixed-radius plots, though BAF and plot size are inversely related. Generally, if variation in the stand is high, BAF should be low; that way variability is "captured" at a sample location and variability between sample locations (which drives sampling error) is made smaller. A good approximation to the relationship between BAF and variability is (Freese 1962, page 27)

$$\frac{CV_{BAF2}^2}{CV_{BAF1}^2} = \sqrt{\frac{BAF2}{BAF1}} \qquad [1\text{-}8]$$

The variation, as measured by coefficient of variation, for 20 BAF is thus about 20 percent greater than that for 10 BAF.

For a fixed cost sampling effort there will be a tradeoff between BAF and number of sample units taken. The smaller the BAF, and hence the more trees tallied at each sample location, the fewer plots that can be taken. A large number of studies have been conducted addressing the issue of optimal BAF. Though study results are varied, it can be shown that, under a reasonable set of assumptions, an optimal BAF (one that results in minimal time/effort to achieve a desired precision) will result in equal amounts of time being spent on travel between plots and measurement at a plot (Zeide 1980). This result holds for random or systematic placement of sample locations, a common practice for timber stand assessments.

Unfortunately, tradition often dictates choice of BAF and this often leads to inefficiencies. A BAF of 10 for Lake States forestry practice is likely too small on average; too much time is spent on an individual plot.

The ability to "see" trees from a sample location should also influence specification of BAF. Large trees can be a long distance from a sample location and still be "in" with a small BAF. Heavy understory conditions also makes ensuring that all "in" trees are sighted difficult. It has been commonly found that using a BAF that is "too small" results in significantly negative measurement bias (too few trees counted). Several studies have shown that it is efficient to tally 5 to 8 trees at a sample location (Wensel *et al*. 1980). To tally 7 trees, on average, at a sample location in a stand estimated to have 120 square feet per acre would imply use of a 17 BAF instrument.

The "5 to 8 tally guideline" has led some to employ a scheme where a different BAF is used at each sample location in a tract. One approach is to start with a large BAF and try successively smaller BAFs until the desired tree count is reached. This is not a good rule to follow, as it can lead to confusion in estimation, though there is some argument as to whether it produces biased results (Iles and Wilson 1988).

An angle gauge is any object of known width held at a known distance from the cruiser's eye. Manipulating equation [1-1] shows that distance from eye (l) and angle gauge width (w), both in inches, are related by

$$l = \frac{w}{2} / \tan\left(\arcsin\sqrt{BAF / 43560}\right) \qquad [1\text{-}9]$$

To use a 1-inch wide gauge for 20 BAF would require holding it 23.33 inches from your eye.

You can also calibrate an angle gauge by backing away from a tree of DBH D (inches) until the tree is just on the plot border and then measuring the horizontal distance R (feet) from the eye to tree center.

$$BAF = 75.625\frac{D^2}{R^2} \qquad [1\text{-}10]$$

This would be a convenient calibration method if you tallied trees using your thumb at arms length as your angle gauge. Many suggest employing such a calibration as it accounts, in some sense, for the ellipticity of trees in your cruise area.

There is little reason to use a BAF that is integer-valued.

How many sample locations are needed?

Sample size is often necessarily set by time or cost constraints. Still, it is good to have an understanding of the relation between number of sample units, population variability between sample units, and sampling error. Study of a particular situation may lead one to conclude that cost constraints lead to an unnecessarily large sample or a sample that is of little value for the purposes to which the derived information is to be put.

For a simple random sample, the number of units required to obtain a sampling error that is P% of the mean, using sample units with variability (between units) CV is[1]

$$n_e = \frac{CV^2}{P^2} \qquad [1\text{-}11]$$

For example, if a sampling error of 10% is acceptable and CV is 50% (based on BAF 10 plots, say) one should establish 25 sample locations. With BAF 20, equation [1-8] suggests a CV of 59.5% giving a sample size of 36.

If the number of units required to obtain a desired sampling error appears out of line, it is likely that effort should be put into stratifying the tract to be sampled; stratification is the most powerful tool available to the sampler. It may also be the case that the specified desired sampling error is unreasonable. Related to this, note that tract size does not appear in the sample size formula. This implies that small tracts with large variability will require what could be considered a ridiculously large number of sample units; or that desired sampling error for small tracts will need to be relaxed. Similar reasoning suggests that precisely estimating summary breakdowns, such as volume of a particular species-size combination, will require large numbers of sample units.

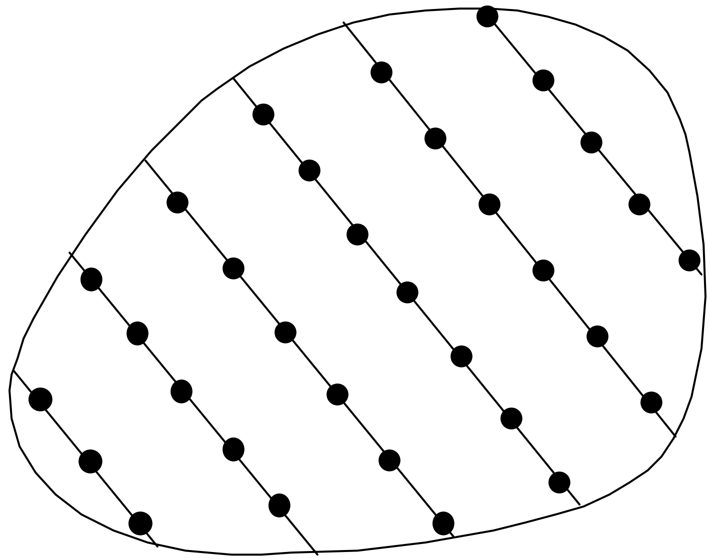
Is it ok to systematically spread out sample locations?

Common sense would suggest placing sample locations on a systematic grid, ala line-plot cruises, throughout the tract of interest in an attempt to obtain as representative a sample as possible. Systematically traversing a tract also expedites mapping, an important task in forest land assessment. Statistical inference, on the other hand, dictates some sort of randomization be used in selecting sample locations. Experience and numerous studies of timber stand assessment applications reach the conclusion that rational application of systematic sampling used in conjunction with standard statistical formulas will produce good point estimates with conservative precision estimates. Rational application means that known trends in the tract, with respect to a variable of interest, must be represented appropriately in the sample. A common guideline is to run cruise lines perpendicular to known trends. Tract edge areas can be particularly difficult to represent properly, a special concern in long, narrow tracts. Aerial photography, existing maps, etc. will usually be sufficient to lead the forester to a rational systematic layout whose results can be summarized using standard statistical formulas.

The figure below illustrates a line-plot cruise with six sample lines and a total of 32 sample locations.

---

[1] Note that the sample size development here differs from many textbooks where a sample size is computed that gives a certain confidence interval half width.



Plots are normally located by compass and pace in timber stand assessments so that pinpointing exact random locations would be impractical. Of greater importance is that sample locations be placed in an unbiased fashion as will be discussed below.

Field implementation

Instrument usage – Angle gauges must be held at the proper distance from the eye, perpendicular to the line of sight, and not tilted along their vertical axis. The cruisers eye must constantly be over the sample location when viewing trees; doing otherwise will result in tally of too many trees. Similar care must be applied when using a prism. The prism itself is the vertex of the critical angle and must therefore always be held over the sample location. A common tendency with a prism is to "push the point" (Oderwald and Gregoire 1995), the bias of which can be significant for large BAFs and small trees.

The line of sight for either instrument is assumed to be horizontal from the sample location to tree breast height. Slight inclinations are not of concern; instruments can be adjusted for slope or a questionable tree can be checked (see below).

Trees hidden by other trees or obscured by brush will likely need to be checked to determine whether they should be tallied. Moving off the sample location should only be done with extreme caution. Some will view the tree above breast height, where brush obstructions are likely less severe, to observe whether the tree is "obviously in"; recall, however, the necessity of proper instrument orientation.

Clearly, missing an entire tree or including a tree that is "off" plot are errors that should and can be avoided with little extra effort.

Checking questionable trees – Difficult conditions and inexperience will make it difficult to judge whether a tree is "in" with an instrument. The DBH (to 0.1-inch) and distance from sample location (horizontal to tree center) must be measured for such trees. DBH of the tree in question should be plugged into the formula

$$R = DBH\sqrt{\frac{75.625}{BAF}} \qquad [1\text{-}12]$$

If the measured distance is greater than the computed distance, the tree should not be tallied. The radical term is often called the plot radius factor and is the number of feet per inch of DBH that a tree can be from the sample location and still be "in."

Some suggest applying the "tally every other one" rule to questionable trees. If interest lies solely in basal area per acre of all trees and several plots are being installed, such a rule will be adequate. In all other situations the rule should be avoided (Iles and Fall 1988). In their study involving experienced cruisers Iles and Fall (1988) found that only about 10 – 15 percent of trees are borderline.

Establishing sample locations in an unbiased fashion – Accurate placement of sample locations is inconsistent with the methods commonly used in timber stand assessments. Of far greater importance is unbiased placement. If the location is inside a tree, so be it; this is of special concern for large BAFs. Overly open or dense areas must not be avoided because they "appear" unrepresentative.

Plots near the tract edge – If the tract of interest contains a moderate to high proportion of edge and edge areas are suspected to differ from interior areas with respect to variables of interest, it is critical that edge areas be appropriately represented in the sampling effort. Plots must not be moved away from tract boundaries. In extreme cases it may be advisable to stratify the tract into edge and non-edge strata and sample the two strata separately. More often, edge conditions can be appropriately treated by taking partial plots or using the mirage method of boundary overlap correction (Avery and Burkhart 1994, page 228). With the mirage method a sample location near the tract edge is tallied twice: once at the normal location and once from outside the tract. The position outside the tract is found by proceeding along the cruise line a distance outside the tract boundary equal to the distance the normal location lies from the tract boundary. Trees "in edge" will be tallied twice with the mirage method. However, the location can be treated like any other single sample location as concerns data summary.

Data summary example

The Excel workbook SinglePhase.xls summarizes data from a variable-radius plot timber stand assessment. Data are entered on the Data worksheet; this includes BAF, number of plots, and for each tree: plot number, tree number, species (user chosen labels, up to seven for one cruise), DBH (integer), and number of 8-foot bolts (integer). Tree volumes in rough cords are computed using Table 6 of Gevorkiantz and Olsen (1955) (worksheet VolTable). Summaries appear on the Results worksheet and include cordwood volume per acre, its standard error, and per acre stock and stand tables. You can also compute the number of plots needed to achieve a specified sampling error expressed as a percent (equation [1-11]). Instructions for using the workbook appear on the Results and Data worksheets.

Double sampling for timber attributes with variable-radius plots

Motivation

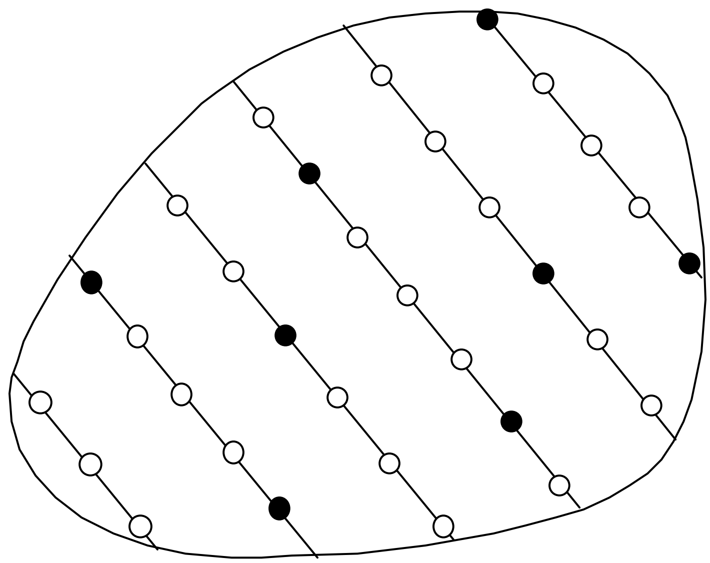
So-called double or two-phase sampling involves the observation of an additional auxiliary variable to improve on the sampling efficiency for estimating some primary variable of interest. To achieve efficiency gains, the auxiliary variable must be closely related to the variable of interest and must be cheaper to observe than the variable of interest. For the case we'll consider, related means linearly related so that the correlation coefficient between the auxiliary variable and the variable of interest will be a convenient measure of relationship closeness.

For timber stand assessments the variable of primary interest will be some measure of volume per acre. A candidate auxiliary variable is therefore basal area per acre. Volume and basal area per acre will typically have a correlation above 0.9 or 0.95. Basal area per acre will be much cheaper to observe than volume when we use variable-radius plots as it requires only a tally of trees "in" at the sample location.

Sampling plan

With double sampling, the auxiliary variable is observed on some large set of plots, n_L, while the primary variable of interest is also observed on a subset, n_S, of the larger set of plots. When using variable-radius plots for a timber stand assessment this means simply tallying trees on the n_L - n_S plots while tree characteristics required to obtain tree volume (DBH, height, …) are measured for "in" trees on the n_S plots. The n_S plots are often referred to as "measure plots" while the n_L – n_S plots can be called "tally plots." Note that the auxiliary variable, basal area per acre, is "free" on the n_S plots.

A double sample can be laid out much like a line-plot cruise. The n_L total plots are placed on a systematic grid in the tract of interest, with attention paid to rational placement of the sample lines. Some fraction (every second, third, fourth, fifth, etc.) of the total number of plots is identified for tree measurement: the n_S plots. The following schematic illustrates for a total of 32 plots where every fourth plot involves tree measurement (solid circles).



Tree tally and measurement

The subset of plots that involve tree measurement, n_S, should be tallied in the same manner as a standard variable-radius plot cruise. Measurements will normally be those necessary to compute individual tree volumes. Tallies must be kept separate by plot if statistical precision estimates are to be computed. For the remaining $n_L - n_S$ plots it is only necessary to accumulate the number of "in" trees across all plots, even if precision estimates are required.

In multi-species stands where estimates of volume per acre by species are of high interest, the cumulative tree tallies on the $n_L - n_S$ plots should be further broken down by species. Similarly, if estimates by tree DBH class are a high priority, one might consider ocularly estimating DBH class for "in" trees on the $n_L - n_S$ plots and completing the cumulative tally by DBH class, or possibly species and DBH class. The extra cost associated with implementing such refinements must be considered and be rationalized to be outweighed by the gains in precision for important estimate breakdowns. This is likely to be the case if experienced cruisers are involved (where species identification and DBH class, 1-inch or 2-inch classes, assignment are automatic).

Volume per acre estimation

Several alternatives are available for summarizing timber stand assessments conducted using double sampling. We will consider the ratio-of-means estimator. The ratio-of-means estimator is appropriate if the variable of interest and auxiliary variable are linearly related with the relationship passing through zero and the variance of the variable of interest is proportional to the magnitude of the auxiliary variable. These assumptions are reasonable for most volume per acre – basal area per acre combinations considered in practice. However, the assumptions should be evaluated and an alternative estimator applied if necessary (cf. Freese 1960, page 39). An example of an instance where the ratio-of-means estimator is inappropriate would be using the auxiliary variable "basal area per acre of <u>all</u> trees" when the variable of interest is board foot volume per acre of trees 10-inches DBH and larger; the intercept of the relationship between these two variables will be significantly negative and the ratio-of-means estimator, consequently, biased.

Equation [1-2] can be applied to the n_S plots upon which tree measurements are taken to obtain a volume per acre value, $\overline{V}_S$. Equation [1-4] can be applied separately to the n_S and n_L plots to get two basal area per acre values, $\overline{G}_S$ and $\overline{G}_L$. The improved estimator for volume per acre is then

$$V_{RMD} = \frac{\overline{V}_S}{\overline{G}_S} \overline{G}_L \qquad [2\text{-}1]$$

The estimator intuitively increases or decreases the volume per acre estimate based on the n_S plots a fraction depending on how the larger sample estimate of basal area per acre compares to the smaller sample estimate.

An estimator for standard error of [2-1] is

$$s_{V_{RMD}} = \left[\left\{ s_V^2 - 2\left(\frac{\overline{V}_S}{\overline{G}_S}\right) s_{VG} + \left(\frac{\overline{V}_S}{\overline{G}_S}\right)^2 s_G^2 \right\} \left(\frac{1}{n_S} - \frac{1}{n_L}\right) + \frac{s_V^2}{n_L} \right]^{1/2} \qquad [2\text{-}2]$$

where s_V^2 is equation [1-5] applied to the n_S volume per acre values and s_G^2 is found using the same equation applied to the n_S basal area per acre values. The n_S <u>pairs</u> of volume per acre – basal area per acre values are used to find the covariance between volume per acre and basal area per acre

$$s_{VG} = \frac{\sum_{j=1}^{n_S} V_j G_j - \frac{\left(\sum_{j=1}^{n_S} V_j\right)\left(\sum_{j=1}^{n_S} G_j\right)}{n_S}}{n_S - 1} \qquad [2\text{-}3]$$

Estimation for categories (breakdowns) of interest

Equation [2-1] can be applied to categories of volume per acre. Volume per acre for a particular species is estimated by finding volume per acre for the species using the measure plots (equation [1-2]) and then multiplying by the ratio $\overline{G}_L / \overline{G}_S$. The same could be done for volume of a species-DBH category and thus an entire stock table (Oderwald 1994). Basal area per acre can be broken down in the same manner.

Applying this approach to stems per acre by species-DBH categories (first using equation [1-3], for example) produces a stand table consistent with the overall estimate of basal area per acre. This is the usual suggestion though it implicitly assumes stems per acre and basal area per acre are linearly related.

What BAF to use?

The same considerations hold in choosing a BAF for double sampling with variable-radius plots as did with standard sampling with variable-radius plots (discussed at length earlier). The 5 to 8 trees per sample location guideline is the simplest to follow. Optimal BAF has yet to be worked out. Normally a single BAF is used for all plots (both measure and tally plots).

What should n_L and n_S be?

Oderwald and Jones (1992) present formulas for choosing n_L and n_S for timber stand assessments using double sampling with the ratio-of-means estimator. They assume specification of the number of standard plots required to meet a precision specification (computed, e.g., using equation [1-11]) and compute n_L and n_S such that the same precision is achieved and cost (time) is minimized. The resulting formulas are complex functions of

- ratio of time to observe volume per acre to time to observe basal area per acre at a sample location
- correlation coefficient between volume per acre and basal area per acre
- ratio of the coefficient of variation of volume per acre to the coefficient of variation of basal area per acre

The worksheet SamSizeP in the Excel workbook TwoPhase.xls implements the formulas. For example if we assume the time for observing volume per acre is four times that of observing basal area per acre, a correlation between volume per acre and basal area per acre of 0.9 (conservative), and that the ratio of coefficients of variation is 1.5, the optimal double sample would measure trees on only one-half the plots of a standard sample but observe a total of 50 percent more plots (n_L) for basal area per acre. Oderwald and Jones cite considerable experience indicating that the ratio of coefficients of variation will be around 1.0 when interest lies in pulpwood-sized timber to less than 1.5 for sawtimber-sized timber. The series of tables in worksheet SamSizeP show how sensitive the results are to typical values of the cost and coefficient of variation ratio; the degree of sensitivity is small to moderate.

The double sampling sample sizes that give minimum variance for a fixed cost (time) are derived in Appendix 1. They are again functions of the three factors identified above. The worksheet SamSizeC in the Excel workbook TwoPhase.xls implements the formulas. For the example above the solution is $n_L = 2\ n_e$ with measurement on every third plot. It is also shown in Appendix 1 that the ratio n_S to n_L is the same for both the equal precision and equal cost sample size solutions.

If data are compiled separately by sample location under existing line-plot cruise procedures, those data can be examined and analyzed to obtain some understanding of the gains possible from double sampling. It should be kept in mind though that the total number of plots tallied under double sampling will typically be larger than that of a standard line-plot cruise.

Another look at the double sampling ratio-of-means estimator

The ratio-of-means estimator we have been considering (equation [2-1]) can be expanded using equations [1-2] and [1-4] as follows

$$\begin{aligned} V_{RMD} &= \frac{\overline{V}_S}{\overline{G}_S}\overline{G}_L \\ &= \frac{\frac{BAF}{n_S}\sum_{j=1}^{n_S}\sum_{i=1}^{t_j}\frac{v_i}{B_i}}{\frac{BAF}{n_S}T_S}\overline{G}_L \\ &= \overline{VBAR}\cdot\overline{G}_L \end{aligned} \qquad [2\text{-}4]$$

where $\overline{VBAR}$ is the average of the individual tree volume to basal area ratios for the T_S trees on the measure plots. The estimator for single phase sampling (equation [1-2]) can be written in this form as well. Volume per acre estimation when sampling with variable-radius plots is thus, generally, seen to be a matter of estimating an average tree VBAR and an average basal area per acre and forming their product.

It is critical to understand that the average tree VBAR in [2-4] (or its single phase equivalent) is obtained with trees selected according to their basal area, i.e. via variable-radius plots. This "average" would, of course, be biased high if interest lies in mean volume to basal area ratio for all trees. Conversely, obtaining the average tree VBAR in [2-4] via equal probability sampling would result in a volume per acre estimate that was biased low.

The relative (to the mean) variances of the two components of a product contribute equally to the relative variance of the product (cf. Freese 1962, p. 17). Individual tree volume to basal area ratios typically exhibit lower variability than plot-to-plot basal area per acre values. Plus the sampling error of average VBAR is controlled by the number of trees measured, a number that is usually larger than the number of sample locations observed. Equation [2-4] clearly points out the advantage of the double sampling approach for timber stand assessments: a greater number of plots are taken observing basal area per acre thus reducing the variability in $\overline{G}_L$; a fewer number of plots, actually fewer measure trees, are required to control the variability in $\overline{VBAR}$.

Equation [2-4] suggests a general approach to volume estimation for double sampling with a ratio-of-means estimator and variable-radius plots (or single phase sampling with variable-radius plots as well): find the best estimate of basal area per acre corresponding to the volume of interest and multiply by the appropriate volume to basal area ratio.

Estimation for categories when extra data are collected on tally plots

It was suggested above that instead of observing only a tree count on the $n_L - n_S$ plots, tree counts by species or species and DBH class (ocularly) might be practical. Such information can be used to improve estimates of volume per acre breakdowns. We can apply the rule gleaned from the previous section: find the best estimate of basal area per acre corresponding to the volume of interest and multiply by the appropriate volume to basal area ratio.

If tree counts are available by species for all plots, our best estimate of basal area per acre for a species comes from all the plots (the total number of trees tallied of that species * BAF / n_L). That quantity should be multiplied by the species average of VBARs from all measure trees of that species. The same approach could be applied if tree counts were available by tree size (DBH) or product class (as examples).

If tree counts are available by species <u>and</u> DBH for all plots, we can obtain a good estimate of basal area per acre from all n_L plots for any species-DBH combination. Multiplying that quantity by the average of VBARs from all measure trees of that species-DBH combination gives the best estimate of volume per acre for that species-DBH combination.

It should be pointed out that the sum of volume per acre estimates across categories obtained in the suggested manner will not be equal to the estimate of volume per acre found using equation [2-1] unless all the category VBAR averages are the same: a simple matter of algebra. However, the difference will normally be small.

The fact that VBARs typically have low variability is important here. We may have few VBAR values for a rare species or for any species-DBH combination. This suggests that if strong interest lies in volume breakdowns we may need to take a larger than optimal (optimal from the viewpoint of estimating total volume per acre) measure sample (n_S plots) to ensure we obtain sufficient VBAR values for estimating breakdowns. The alternative is to “borrow” VBAR information from “similar” species or sizes of trees. It should be pointed out that VBARs will vary most as a function of tree height if we consider only merchantable size, commercially-important Lake States tree species.

Species estimates of basal area per acre can be used to construct improved stand table estimates as well. For a particular species, usually defining a column of a stand table, the ratio $\overline{G}_L / \overline{G}_S$ <u>computed for the species</u> should be multiplied by each estimate of trees per acre (obtained from the measure plots only). The example below illustrates for a species where basal area per acre was 17.58 on the measure plots but 22 for all plots. If both species and DBH class (ocularly) are observed on tally plots, a stand table can be constructed directly from those data;

this assumes ocular DBH estimates have the same information as measure plot DBH measurements.

DBH	Stems per acre (unadjusted)	Basal area per acre (unadj.)	Stems per acre (adjusted)	Basal area per acre (adjusted)
5	22	3.00	27.53	3.75
6	32	6.28	40.04	7.86
7	18	4.81	22.52	6.02
8	10	3.49	12.51	4.37
	Total	17.58		22

Tally sheets

A tally form for double sampling with variable-radius plots should track tree tally (possibly by species) for the $n_L - n_S$ plots and tree measurement (species, DBH, height, …) on the n_S plots. If sampling error is to be computed, separate sheets are required for each of the n_S plots.

Data summary example

The Excel workbook TwoPhase.xls summarizes data from a double sample with variable-radius plots timber stand assessment using the ratio-of-means estimator. Data are entered on the Data worksheet; this includes BAF, n_L, n_S, number of tally trees by species across ALL (n_L) plots and for each measure tree: plot number, tree number, species (user chosen labels, up to seven for one cruise), DBH (integer), and number of 8-foot bolts (integer). Tree volumes in rough cords are computed using Table 6 of Gevorkiantz and Olsen (1955) (worksheet VolTable). Summaries appear on the Results worksheet and include cordwood volume per acre and its standard error and per acre stock and stand tables. Instructions for using the workbook appear on the Results and Data worksheets. This workbook is not as automatic as SinglePhase.xls, as you are required to manually produce the stock and stand tables that are improved using the species basal area data assumed to be collected on all plots.

The Big BAF method

Motivation

Formulation [2-4], which we have seen is general for volume per acre estimation with variable-radius plots, and its associated relative variance suggest the average VBAR and basal area per acre estimates that constitute the variable-radius plot volume per acre estimator need not be linked. One should strive to obtain as good an estimate (low variance) of each as is possible. In double sampling, both average VBAR and basal area per acre are obtained from the same set of plots. Trees measured for volume, and hence those that will contribute to an average VBAR, are obtained on a subset of the total number of plots, and hence a subset of the total tract area: they are clustered. Number of measure trees drives the sampling error of average VBAR. In double sampling with variable-radius plots the number of measure trees is not directly controlled but rather dictated by the BAF applied to all plots. It is likely that this number is sub-optimal and many forest sampling experts would suggest it is generally too high: often way too high.

So, how do we spread the measure trees around the tract, intuitively appealing, and control their number, statistically appealing? One method is to use a second, possibly much larger, BAF at every sample location to select measure trees (recall that the VBAR trees must be selected in proportion to their basal areas). This method was suggested about 15 years ago by Dr. Kim Iles (personal communication) and has come to be known as the Big BAF method. The method is being successfully applied in the western United States and Canada (cf. guest articles by Corin and Crowther at http://www.proaxis.com/~johnbell/newsindex/guest.htm). There seems no reason why the method would not work equally well in Lake States timber stand assessments.

Sampling plan

The Big BAF method requires specification of a number of sample locations, n_{BBAF}, and two BAFs. The n_{BBAF} sample locations will normally be laid out in a line-plot cruise fashion presumably using the same number of sample lines as an equivalent standard cruise. At each sample location, a tally of trees is made using a standard, labeled smaller, BAF (denoted BAF_S), and measure trees, those for which VBAR will be determined, are selected using a second, larger BAF (denoted BAF_B). Just as suggested with double sampling, tally trees may have species and ocular DBHs recorded as well.

Volume per acre estimation

A simple average VBAR can be computed using the measure trees selected with BAF_B. For clarity we will denote that $\overline{VBAR}_{BBAF}$. The best estimate of basal area per acre is obtained using the n_{BBAF} tallies based on BAF_S (only): again equation [1-4]. We will denote that $\overline{G}_{BBAF}$. Trees tallied with BAF_B will, of course, be tallied with BAF_S as well. Taking the product of the two estimators gives the Big BAF estimate of volume per acre

$$V_{BBAF} = \overline{VBAR}_{BBAF} \cdot \overline{G}_{BBAF} \quad [3\text{-}1]$$

Note that, applying the reverse of the logic used to arrive at equation [2-4], $\overline{VBAR}_{BBAF}$ is equivalent to the ratio of average volume per acre to average basal area per acre using standard formulas [1-2] and [1-4] based on the BAF_B tallies.

Using SE%(X) to denote standard error of the estimator X as a percent of the estimator X, the suggested estimator for standard error of [3-1] is

$$SE\%(V_{BBAF}) = \sqrt{SE\%(\overline{VBAR}_{BBAF})^2 + SE\%(\overline{G}_{BBAF})^2} \quad [3\text{-}2]$$

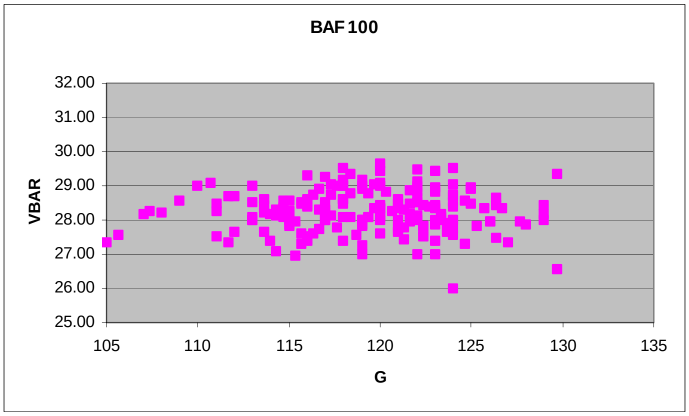
where

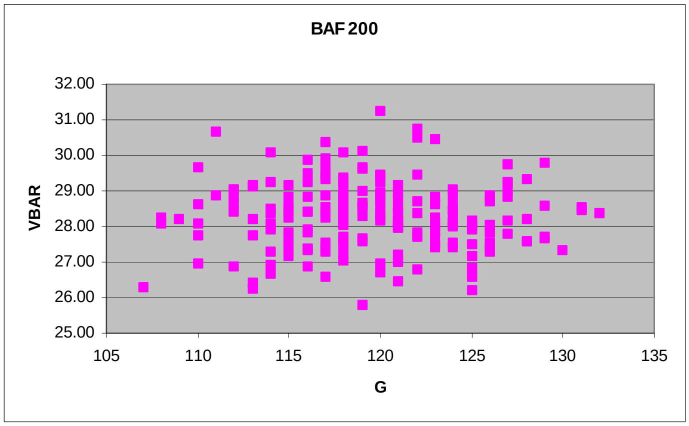
$$SE\%(\overline{VBAR}_{BBAF}) = \sqrt{\frac{\sum_{k=1}^{T}\left(\frac{v_k}{B_k}\right)^2 - \frac{\left(\sum_{k=1}^{T}\frac{v_k}{B_k}\right)^2}{T}}{T(T-1)\overline{VBAR}_{BBAF}^2}} \cdot 100$$

$$SE\%(\overline{G}_{BBAF}) = \sqrt{\frac{\sum_{j=1}^{n_{BBAF}} G_j^2 - \frac{\left(\sum_{j=1}^{n_{BBAF}} G_j\right)^2}{n_{BBAF}}}{n_{BBAF}(n_{BBAF}-1)\overline{G}_{BBAF}^2}} \cdot 100$$

and T is the total number of measure trees across the n_{BBAF} sample locations using BAF_B and G_j is basal acre per acre at the jth sample location tallied using BAF_S. The standard error for $\overline{VBAR}_{BBAF}$ assumes the tree VBARs are i.i.d. which will be reasonable for large BAF_Bs.

Formulation [3-2] has been found to give reasonable results in practice (Iles, personal communication), giving conservative sampling error estimates (over-estimates) if indeed the two components of the product are independent. Current research suggests the independence assumption is reasonable (cf. figures below).





Number of sample locations, number of measure trees, and implied BAF_B

The total number of sample locations, $n_{BBAF,}$ and measure trees, $T_{BBAF,}$ required to achieve, at minimum cost, a sampling error equivalent to the simple random sample n_e can be found by applying the approach of Oderwald and Jones (1992). The derivation is presented in Appendix 2. The resulting formulas are not intuitive and involve the cost ratio

$$\frac{c_S + c_B}{c_m}$$

where

$$c_S = \text{cost of tallying a plot with BAF}_S$$
$$c_B = \text{cost of tallying a plot with BAF}_B$$
$$c_m = \text{cost of measuring a tree once it is determined "in"}$$

These costs (times) are best expressed in terms of expected number of measure trees per plot with BAF_S, BAF_B, and

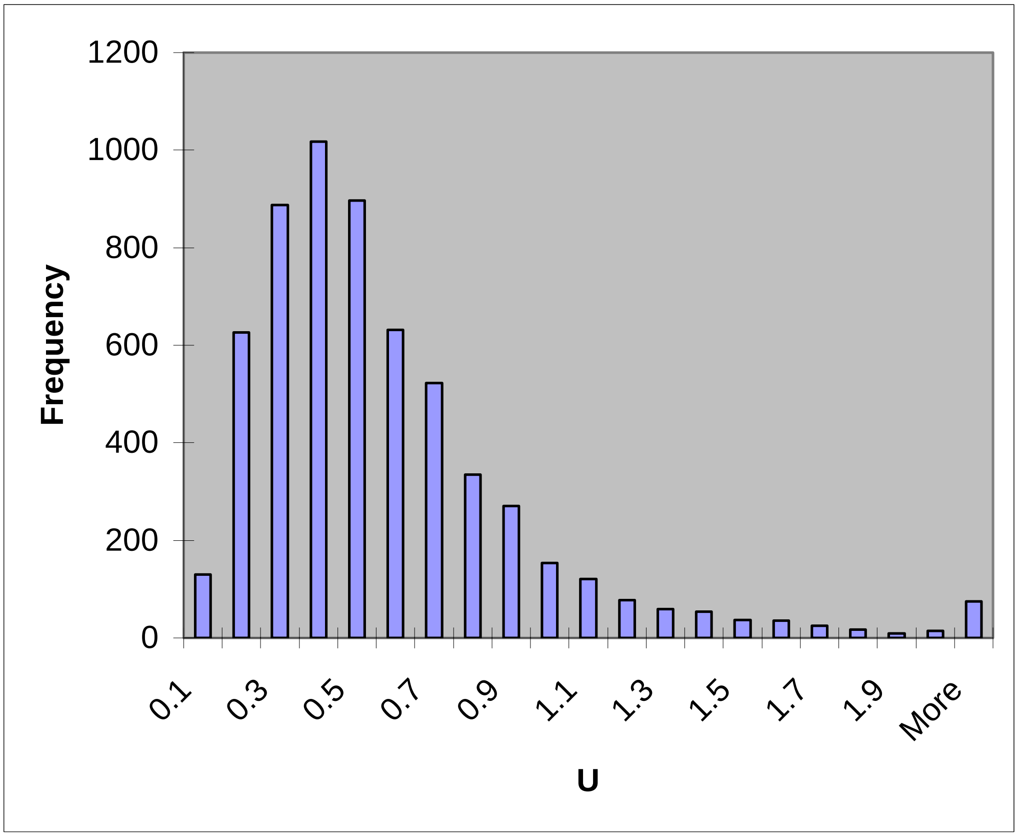
$$k = \frac{\text{cost of measuring trees identified to be "in" at a location}}{\text{cost of tallying trees at a location}}$$

The solutions for n_{BBAF} and T_{BBAF} are implemented in worksheet SamSizeP of workbook BigBAF.xls. To use the worksheet requires specifying n_e, k, BAF_S, expected basal area per acre, and

$$B = \frac{CV_V}{CV_G}$$

$$U = \frac{CV_{VBAR}}{CV_G}$$

Typical values for k are {2 – 5} depending on the tree measurements taken. B was introduced with double sampling where it was suggested that values ranging from 1 to 1.5 are common; it is required here that the two CVs in B are determined with the same BAF (BAF_S). BAF_S will be chosen using the 5 to 8 trees per sample location guideline. Computation of U from nearly 6000 stand cruises resulted in the following frequency distribution.



The mean value of U was approximately 0.5, though its distribution is quite skewed to the right. Still, nearly 75% of the values of U were between 0.2 and 0.7, while almost 90% of the U values were less than 1.0. Since measure trees are selected with probability proportional to tree basal area, variability in tree VBAR closely parallels variability in tree height. Stands where U exceeds 1.0 are therefore those where the (relative) variability in tree height exceeds the spatial variability in basal area per acre.

An iterative solution is required for n_{BBAF} and T_{BBAF} with a guess for BAF_B used as a starting point. The Excel add-in Solver can be used. The final BAF_B is optimal for the supplied inputs and consistent with the solution for the optimal n_{BBAF}, T_{BBAF} pair.

It is unlikely that you will be able to purchase a prism or angle gauge corresponding to an optimal BAF_B. A good solution is to construct a rectangular angle gauge whose short side width corresponds to BAF_S and whose long side width corresponds to BAF_B, with the instrument held a comfortable distance from the eye. Equation [1-9] can be used to determine the dimensions of the rectangle and is coded into worksheet SamSizeP in Excel workbook BigBAF.xls. Exacting construction of the angle gauge is not required as the value of BAF_B itself is not used in any calculation.

Since T_{BBAF} is expected total number of measure trees, it is possible that the actual total number of measure trees, upon completion of the cruise, will be too low. BAF_B can be decreased somewhat to help ensure this does not happen; results will be suboptimal in terms of minimizing cost. Also see related discussion below.

Values for n_{BBAF} and T_{BBAF} that result in minimum variance for a cost equal to simple random sampling are also derived in Appendix 2 and are implemented in worksheet SamSizeC of workbook BigBAF.xls.

The optimal results for number of tally plots and measure trees (and implicitly BAF_B) for Big BAF sampling will seldom be strictly applied. The sample size worksheets in workbook BigBAF.xls should be experimented with using conditions of interest. It is likely that a set of four or five BAF_B angle gauges (perhaps 80, 120, 160, 200, and 240 for Lake States forestry practice) will adequately cover the range of circumstances a cruiser will encounter.

Additional comments

Proper tree tally – When BAF_B is large, as it will be under circumstances where the Big BAF method works best, it is important that all the details of proper instrument usage are given appropriate attention. The instrument must be held at the proper location and in proper orientation. Questionable trees must be carefully checked. The effort involved is lessened by the fact that questionable trees will be close to the sample location. Since BAF_B does not explicitly appear in any calculation formula it is only important that measure trees are consistently selected, regardless of conditions. Equation [1-12] has been coded into worksheet SamSizeP in Excel workbook BigBAF.xls to aid in checking questionable trees. Sample locations must be placed in an unbiased fashion, inside of trees if necessary, and edge conditions must be properly accounted for.

It might seem that use of BAF_B for measure trees would result in only "large" trees being measured for VBAR. Equation [1-1] clearly shows this is not the case. The odds of measuring a tree whose basal area is x-times that of another tree is x:1, independent of BAF.

Sufficient numbers of measure trees – As mentioned with double sampling, if categorizations (breakdowns) of volume per acre of interest, e.g. volume per acre by species or a stock table, it is necessary to have an estimate of average VBAR for each category (or to "borrow" VBAR information). If the expected number of measure trees, T_{BBAF}, is too small to adequately fill all categories, it could be adjusted upwards and a new, sub-optimal in terms of expected costs, n_{BBAF}

computed. Worksheet SamSizeP of workbook BigBAF.xls allows this. Another suggestion is to measure representative trees of rare categories as they are encountered, though caution should be exercised to avoid selection bias. If a particularly rare category is of high importance a separate, smaller, BAF might best be used for selecting measure trees for the category.

Estimation for categorizations of volume per acre, as well as other attributes of interest, can proceed as outlined with double sampling, including use of extra data collected on tally plots. Volume per acre for a category is again the product of basal area per acre for the category multiplied by average VBAR for the category. Volume per acre categories (volume for a species and product class for example) will typically exhibit lower tree-to-tree variability in VBAR than what will be found across categories. However, plot-to-plot variability in basal area per acre will likely be greater for a category than for all categories combined. This again emphasizes the importance of tallying more plots (getting more observations of basal area per acre), an outcome the Big BAF method promotes. Still, most cruisers would be uneasy if no trees were measured for a particular category of interest, requiring the borrowing of VBAR information from another category. Ideally, two to five measure trees per category should be obtained.

The Big BAF advantage – The Big BAF method trades off tree measurement for tree tally. Focusing more of the sampling effort on tree tally at a larger (than simple random sampling) number of sample locations results in improved basal area per acre estimates, the major source of variability in estimating volume per acre. If one only reduces tree measurement intensity by selecting measure trees with a larger BAF, the advantage of the Big BAF method will not be realized.

Tally sheets

A tally form for Big BAF sampling should track tree tally (possibly by species) for the n_{BBAF} plots and tree measurement (species, DBH, height, …) of trees selected with BAF_B. If sampling error is to be computed, separate tree tallies are required for each of the n_{BBAF} plots.

Data summary example

The Excel workbook BigBAF.xls summarizes data from a timber stand assessment using variable-radius plots and the Big BAF method. Data are entered on the Data worksheet; this includes BAF_S, BAF_B, n_{BBAF}, number of BAF_S tally trees by species for each plot, and for each measure tree: tree number, species (user chosen labels, up to seven for one cruise), DBH (integer), and number of 8-foot bolts (integer). Tree volumes in rough cords are computed using Table 6 of Gevorkiantz and Olsen (1955) (worksheet VolTable). Summaries appear on the Results worksheet and include cordwood volume per acre and its standard error and per acre stock and stand tables. Instructions for using the workbook appear on the Results and Data worksheets. BigBAF.xls is used much like TwoPhase.xls; you are required to manually produce the final stock and stand tables.

A cruise exercise: Conduct and review

A test stand has been laid out for participants to gain experience using the double sampling and Big BAF methods discussed above. The instructor has obtained data from a cruise of the stand using a _____ BAF instrument, measuring all tally trees on _____ plots. The results obtained were

Item	Description	Value
1.	Volume per acre (cords) (V)	
2.	Standard error of estimate (%)	
3.	Basal area per acre (G)	
4.	Average tree VBAR	
5.	CV_V	
6.	CV_G	
7.	CV_{VBAR} (trees)	
8.	Number of trees measured	
9.	Correlation (G, V)	

Each crew of two individuals should run sample lines through the test stand taking plots as follows

a. tally "in" trees at every sample location using a _____ BAF
b. measure "in" trees at every _____ sample location using a _____ BAF
c. measure trees that are "in" with a _____ BAF at <u>every</u> location

Use the sheets provided by the instructor to record the results. "Measuring" a tree means obtaining its species, DBH, and number of 8-foot bolts to a pulpwood top. Be sure to check questionable trees.

Also attempt to collect approximate average tally/measure times as follows

Item	Description	Value
1.	Time to tally trees using _____ BAF	
2.	Time to tally (identify "in") trees using _____ BAF	
3.	Time to tally and measure trees using _____ BAF	

What should the design elements of a double sample be for this stand?

What should the design elements of a Big BAF sample be for this stand?

Role of volume tables/equations

Assuming interest lies in volume per acre estimates, direct measurements of tree dimension will need to be converted to appropriate volume estimates. A wide variety of volume tables have been developed for this purpose. Volume tables differ in terms of inputs required, output units, merchantability standards, and the underlying functional relationship between inputs and outputs. Choice of an appropriate volume table is important as the values obtained from it are most often assumed to be without error from a statistical sampling error calculation perspective.

For timber stand assessment applications in Lake States forestry practice, the volume tables of Gevorkiantz and Olsen (1955) have found wide acceptance; for example they are the basis for the popular "cumulative volume tally sheet." The tables are not explicitly species-specific, though adjustment factors based on species and stand characteristics are provided for circumstances that can justify the extra expense involved in their application. Table 6 of Gevorkiantz and Olsen presents estimates of rough cords as a function of DBH and number of 8-foot bolts. Table 6's popularity is evidenced by several attempts at "formulating" its implied functional relationship; examples are Stott (1962) and Stone (as reported by Hahn 1984). However, Ek and Droessler (1988) have pointed out that caution should be exercised in applying Table 6 as it implies the use of variable top diameters that may be incompatible with present-day merchantability standards. Still, the taper table (Table 8) in Gevorkiantz and Olsen, upon which all other tables are based, remains valid. The taper table can be used to generate a volume table of any merchantability specifications. Burk and Ek (1999) presented an approach to using the taper table in automated processing systems and thoroughly analyzed the tables in Gevorkiantz and Olsen.

Constant form factor volume equations

A constant form factor volume equation relates some measure of tree volume to tree DBH (D) and some measure of tree height (H) as follows

$$v = f \cdot D^2 \cdot H \qquad [5\text{-}1]$$

The appropriateness of using a volume equation of the form [5-1] depends on the particular volume measure and particular height measure being used. If a volume equation of the form [5-1] is being used, volume per acre estimation using variable-radius plots requires <u>measurement of height only</u>.

Substituting [5-1] into [1-2] gives

$$\begin{aligned} \text{volume per acre} &= \frac{BAF}{n}\sum_{j=1}^{n}\sum_{i=1}^{t_j}\frac{fD_i^2 H_i}{0.005454 D_i^2} \\ &= \frac{f \cdot BAF}{0.005454 \cdot n}\sum_{j=1}^{n}\sum_{i=1}^{t_j} H_i \end{aligned} \qquad [5\text{-}2]$$

Table 3 of Gevorkiantz and Olsen (1955) reports total inside bark volume in cubic feet as a function of DBH and total height. The table (for trees greater than 30 feet tall) is equivalent to a constant form factor volume equation with f = 0.00229 . If using a 20 BAF instrument, equation [5-2] with f from Table 3 implies each foot of height represents 8.4 cubic feet per acre at a single sample location.

It is important to understand that tallying by height only based on the above formulation is not an approximation. Under the stated circumstances, measuring DBH adds no information to observing volume per acre. Still, there may be other considerations, e.g. generation of a stock table, that may dictate that DBH be measured/observed.

Volume table variation in VBARs

Equation [5-1] implies no variation in VBAR as a function of DBH. Even though [5-1] may not hold exactly, the variation in VBAR across DBHs may be insignificant relative to the precision requirements of a timber stand assessment. The following table is a portion of Table 6 (Gevorkiantz and Olsen, 1955) with VBAR rather than volume tabulated.

Tree DBH (in.)	No. of 8-foot bolts						
	1	2	3	4	5	6	7
5	0.081	0.140	-	-	-	-	-
6	0.087	0.143	0.204	-	-	-	-
7	0.086	0.142	0.198	0.255	-	-	-
8	0.089	0.143	0.195	0.249	0.304	-	-
9	0.090	0.147	0.199	0.247	0.294	0.346	-
10	0.090	0.150	0.204	0.244	0.294	0.345	0.387
11	0.091	0.152	0.208	0.250	0.288	0.335	0.379
12	-	0.154	0.210	0.252	0.287	0.331	0.382
13	-	-	0.214	0.256	0.291	0.331	0.380
14	-	-	-	0.255	0.291	0.330	0.374
15	-	-	-	-	0.297	0.334	0.375
Average	0.088	0.146	0.204	0.251	0.293	0.336	0.380

For a given number of bolts, VBAR is seen to vary little across DBH. The last row of the table presents average VBAR for each bolt class. These values could be used in a tally by height only where numbers of trees by bolt class were tracked. The amount of error incurred using this approximation is illustrated via an example in the following table. The data are results from tallying/measuring two 20 BAF variable-radius plots.

DBH	Bolts	Tallies	Σ VBAR (approx.)	Σ VBAR (actual)	DBH	Bolts	Tallies	Σ VBAR (approx.)	Σ VBAR (actual)
6	2	1	.146	.143	9	4	2	.502	.494
7	1	1	.088	.086	9	5	2	.586	.588
7	2	2	.292	.284	9	6	1	.336	.346
8	3	1	.204	.195	12	6	1	.336	.331
8	4	2	.502	.498	14	5	1	.293	.291

Volume per acre using the average VBARs is 32.85 cords while volume per acre using the actual VBARs is 32.56 cords, a difference of less than one percent. The accuracy of using VBAR averages depends on the distribution of the tally trees with respect to the table. The VBAR averages can be made more precise for particular circumstances if knowledge concerning the range of expected DBH/No. bolt combinations is known.

The origin of "shortcut formulas"

The shortcut "number of sticks plus number of trees divided by two" has been commonly used for estimating cords per acre when cruising timber with a 10 BAF instrument in the Lake States (Ek and Burk 1986). The origin of this shortcut can be found by analyzing the VBAR table above. Relating average VBAR to number of bolts (via simple linear regression e.g.) gives approximately

$$\text{Average VBAR} = 0.05(1 + \text{No. Bolts})$$

Substituting this into equation [1-2] for <u>one sample location</u> (n=1) gives

$$\text{cords per acre} = BAF \sum_{i=1}^{t} 0.05(1 + \text{No. Bolts}_i)$$

$$= BAF \frac{(t + \sum_{i=1}^{t} \text{No. Bolts}_i)}{20}$$

which for BAF = 10 produces the shortcut. Note that for BAF = 20 the shortcut is simply "number of sticks plus number of trees."

For the above example the shortcut gives (14 + 53)/2 = 33.5 cords per acre, in error by a little less than three percent. The shortcut would seem adequately accurate for "eyeball cruises."

Similar shortcuts can be derived for many volume tables. For example, an analysis similar to the one just presented applied to Table 2 of Gevorkiantz and Olsen (1955) suggests estimating International ¼-inch rule board foot volume per acre at a sample location as BAF * 25 * (No. trees + 2 * No. logs). Ashely (1980) provides several other shortcuts applicable to Lake States forestry practice.

Eliminating tree measurement from timber stand assessments

Equation [2-4] shows that if one wishes to only tally trees (i.e. estimate average basal area per acre), mean VBAR of trees selected via an angle gauge must be known. While it is unlikely that mean VBAR will be known, accumulated experience and some knowledge of stand characteristics may lead to a sufficiently accurate, for the purposes at hand, estimate of mean VBAR.

The mean VBAR of trees selected with an angle gauge is equal to the VBAR of the tree of average VBAR. For many circumstances this can be adequately approximated by the VBAR of the tree of average DBH among trees selected with an angle gauge. This average DBH along with a typical height for a tree of that DBH can be used to approximate mean VBAR.

The probability of selecting a DBH with an angle gauge is proportional to the product of the frequency of trees of that DBH and DBH^2. The average DBH of trees selected with an angle gauge can be shown to be the ratio of the third to second non-central moment of DBH. This DBH will be several tenths of an inch larger than quadratic mean DBH (the DBH of the tree of mean basal area) for a typical pulpwood-sized stand in the Lake States.

In the above example, if we had assumed the average tally tree would have DBH = 9 and that a 9-inch DBH tree typically had 4 bolts, we would estimate an average VBAR of tallied trees of 0.247 giving a volume per acre of 34.6 cords per acre (with the tally of 14 trees at two 20 BAF sample locations constituting our cruise).

If a volume equation of form [5-1] is appropriate for trees in the stand being cruised, then average VBAR can be shown to be

$$\frac{f}{0.005454}\overline{H}$$

where $\overline{H}$ is the average height of trees tallied with an angle gauge. This also suggests using the VBAR of the tree of average height, among those that would be selected with an angle gauge, as our approximation of average VBAR.

Sampling error estimation is, of course, problematic if one adopts any of these *ad hoc* methods.

Local (localized) volume tables

In an individual stand, total tree height is often tightly coupled with tree DBH. This fact can be taken advantage of by converting a volume table/equation based on DBH and total height, to one based on DBH only; a so-called local volume table. Rather than observing the total height of every measure tree, a subsample is instead observed to establish the relationship between total height and DBH. DBH is observed for every measure tree. Once established, the relationship can be used to infer height or volume for trees on which only DBH is measured. Trees for height measurement need not be selected at random; obtaining heights for a broad range of DBHs is of greater importance.

While the approach described can conceptually be applied for volume tables based upon merchantable height, the relationship between merchantable height and DBH is not as strong as that between total height and DBH.

A local volume table/equation can be derived for the described procedure in at least two ways. Total height/DBH data can be related via an equation and the estimated equation substituted for total height in the volume table. A regression equation of the form $\ln(H) = b_0 + b_1/DBH$ is commonly used for this purpose. Alternatively, the observed total height/DBH pairs can be used with the standard volume table to estimate volume for the corresponding trees and subsequently volume predicted directly using an equation like $v = b_0 + b_1DBH^2$ fit to those data. The latter approach will be preferred if it is found to work adequately. Note that if it is possible to describe the volume/DBH relationship with the equation $v = c_1DBH^2$, no tree measurement, beyond the subsample used to observe height and DBH, is required to estimate volume per acre. Unfortunately, $v = c_1DBH^2$ is seldom an adequate equation for relating volume and DBH.

Localized volume equations derived using the approach detailed will most likely be of lower accuracy than volume equations based on height and DBH. The importance of the loss in accuracy must be judged on a case-by-case basis. Local volume equations are popular due to the cost/time involved in height measurement, which they greatly reduce. Part of the apparent accuracy loss is nullified by the difficulty in accurately measuring tree heights without expending considerable effort.

A look at sampling efficiencies

Variable- versus fixed-radius plots

We have seen that various implementations of sampling with variable-radius plots share the estimation formula

$$\text{volume per acre} = \overline{VBAR} \cdot \overline{G} \qquad [6\text{-}1]$$

where $\overline{VBAR}$ is the average volume to basal area ratio of trees selected with an angle gauge and $\overline{G}$ is average basal area per acre. For fixed-radius plots the analogous formula is

$$\text{volume per acre} = \overline{v} \cdot \overline{T} \qquad [6\text{-}2]$$

where $\overline{v}$ is average volume per tree and $\overline{T}$ is average trees per acre.

Given the nature of variable- and fixed-radius plots, $\overline{G}$ and $\overline{T}$ will be estimated with similar relative precision with the respective plot types (for "similar size" plots). However, we have seen that volume to basal area ratios exhibit less tree-to-tree variability than do tree volumes. Since the relative variation in volume per acre is, to a first order approximation, an additive function of the two constituent relative variabilities, variable-radius plots would appear to give more precise estimates of volume per acre than fixed-radius plots.

It should be noted that this is merely an additional intuitive argument in favor of variable-radius plots. Variance of the volume per acre estimator will also be impacted by the sign and magnitude of the correlation between the two constituent parts of [6-1] and [6-2]. This correlation depends on the specifics of the implementation.

Double sampling with the ratio-of-means estimator

Recall Oderwald and Jones (1992) presented formulas for n_L and n_S in double sampling with a ratio-of-means estimator and variable-radius plots that were the <u>minimum cost</u> solution for attaining a precision equal to that of a given simple random sample. They also show that the double sample will have <u>lower cost</u> when

$$K > \frac{B^2}{2\rho B - 1} \qquad [6\text{-}3]$$

where

- K = ratio of time to observe volume per acre to time to observe basal area per acre at a sample location
- ρ = correlation coefficient between volume per acre and basal area per acre
- B = ratio of the coefficient of variation of volume per acre to the coefficient of variation of basal area per acre

The values of K for which double sampling will result in the same cost as simple random sampling are displayed for several combinations of ρ and B in the table below. A larger value of K would mean that double sampling was more time efficient than simple random sampling.

ρ	B = 1	B = 1.5
0.95	1.11	1.22
0.90	1.25	1.32
0.80	1.67	1.61
0.65	3.33	2.36

Using the results from Appendix 1 it can be shown that [6-3] also gives the values of K for which the equal cost solution (cost of double sampling equal to the cost of simple random sampling) results in greater precision (lower variance) than simple random sampling.

Tree measurement does not need to involve much extra cost for double sampling to show significant efficiency (cost or precision) gains for common timber stand assessment circumstances. To determine whether the magnitude of the gain eclipses the inertia/cost involved in implementing a new system requires consideration of organizational issues. However, there seems little doubt that a substantial, long term gain will be realized when applying double sampling with variable-radius plots to the timber stand assessment problem where primary interest lies in estimating volume per acre.

To take a specific example, suppose 25 plots, arranged as a simple random sample, are required to satisfy a specified precision requirement. Suppose further that ρ = 0.925, B = 1, and K = 3. The simple random sample would cost 75 "cost units." The minimum cost double sample solution (using worksheet SamSizeP in workbook TwoPhase.xls) has $n_L = 34$ and $n_S = 11$ which would cost 56 "cost units," a savings of 25 percent. If a "cost unit" has value $7.50, say, the savings would be $142.50 for this one assessment. Alternatively, a double sample with cost equal to the simple random sample (using worksheet SamSizeC in workbook TwoPhase.xls) would suggest taking $n_L = 47$ and $n_S = 14$ with a precision gain of 15 percent in terms of standard error.

The Big BAF method

Cost efficiency comparisons for the Big BAF method are made difficult by the complexity of the formulas involved. The necessary results are given in Appendix 2 and can be implemented as part of the SamSizeP worksheet of workbook BigBAF.xls. It is again instructive to compute the value of k for which equally precise Big BAF and simple random samples cost the same amount. Here

$$k = \frac{\text{cost of measuring trees identified as "in" at a location}}{\text{cost of tallying trees at a location}}$$

Note k = K – 1 where K was used in double sampling. Some results for B = 1 and $n_e = 10$ are

U	Trees tallied per plot with BAF_S	
	5	9
0.3	0.63	0.46
0.5	1.10	0.76
0.7	1.84	1.15
0.9	3.20	1.71
1.1	5.97	2.50

Recall that B will normally be 1.0 or larger. For larger values of B, the tabled values become smaller, often significantly. For larger values of n_e, the tabled values also become smaller, though to a very small degree. The table would indicate that, under most common circumstances, there is a clear and substantive advantage in using the Big BAF method.

The same analysis can be applied to the case of fixed cost, also derived in Appendix 2. Using worksheet SamSizeC of workbook BigBAF.xls it can be shown that the above tabled values of k are also those for which the equal cost solution gives precision equal to simple random sampling.

To continue the example from the previous section assuming k = 2, U = 0.5, and that BAF_S will be chosen to average 7 tally trees per plot, we find that $n_{BBAF} = 34$ and $T_{BBAF} = 27$ will cost 70% of simple random sampling but give the same precision. The source of this gain is largely due to going from measuring 175 trees with simple random sampling to 77 trees with double sampling to 27 trees with the Big BAF method. For a cost equal to simple random sampling and the same conditions, the optimal Big BAF sample would have $n_{BBAF} = 48$ and $T_{BBAF} = 38$ and result in a precision gain of 16 percent in terms of standard error.

Summary

The gains from using double or Big BAF sampling are clear and often significant. In most instances Big BAF is the superior method for estimating a single volume per acre value; this is not surprising given it is the only method that accounts for the disparity in variability between the two constituent parts of volume per acre and explicitly selects a number of trees to measure. Caution must only be exercised in circumstances where the variability in VBARs is large relative to the variability in basal area per acre and few trees are tallied with BAF_S; such circumstances are rare.

If volume per acre is only one of several attributes of interest, the advantage of double and Big BAF sampling will not be as straightforward. In particular if categorizations of volume per acre are of interest it is important that sufficient tree numbers are measured to estimate VBAR for the various classes of interest. Equally or more important though is that basal area per acre is well estimated for the classes and Big BAF sampling addresses that. In fact, as classes become more specific, the variability in tree VBAR within a class will likely diminish significantly while the plot to plot variability in class basal area per acre will likely increase significantly; this clearly suggests a division of sampling effort consistent with the Big BAF method.

References

Ashely, B.S. 1980. Short-cut inventory procedures can be accurate and useful. Proceedings forest inventory for private nonindustrial woodlands, eds. B.C. Fisher and H.A. Holt. 33-39. J.S. Wright Forestry Conference, Department of Forestry and Natural Resources, Indiana Cooperative Extension Service, Purdue University.

Avery, T.E. and H.E. Burkhart. 1994. Forest measurements. Fourth Edition. McGraw-Hill, Inc. New York. 408 pp.

Burk, T.E. and A.R. Ek. 1999. Approximation, evaluation, and extension of the Lake States composite taper and volume tables. Northern Journal of Applied Forestry 16:65-71.

Ek, A.R. and T.E. Burk. 1986. Why number of sticks plus number of trees divided by two may equal cords per acre. Northern Journal of Applied Forestry 3:133-134.

Ek, A.R. and T.D. Droessler. 1986. An error in application of the composite volume table: Table 6 of USDA Technical Bulletin 1104. Northern Journal of Applied Forestry 3:88.

Freese, F. 1962. Elementary forest sampling. USDA Handbook No. 232. 91 pp.

Gevorkiantz, S.R. and L.P. Olsen. 1955. Composite volume tables for timber and their application in the Lake States. USDA Technical Bulletin No. 1104. 51 pp.

Goodman, L.A. 1960. On the exact variance of products. Journal of the American Statistical Association 55:708-713.

Hahn, J.T. 1984. Tree volume and biomass equations for the Lake States. USDA Forest Service Research Paper NC-250. 10 pp.

Iles, K. and M. Fall. 1988. Can an angle gauge really evaluate “borderline trees” accurately in variable plot sampling? Canadian Journal of Forest Research 18:774-781.

Iles, K. and W.H. Wilson. 1988. Changing angle gauges in variable plot sampling: Is there a bias under ordinary conditions? Canadian Journal of Forest Research 22:768-773.

Oderwald, R.G. 1994. Stock and stand tables for point, double sampling with a ratio of means estimator. Canadian Journal of Forest Research 24:2350-2352.

Oderwald, R.G. and T.G. Gregoire. 1995. Overstated volumes from “pushing” the point. Southern Journal of Applied Forestry 19:162-165.

Oderwald, R.G. and E. Jones. 1992. Sample sizes for point, double sampling. Canadian Journal of Forest Research 22:980-983.

Stott, C. 1962. The new large computers require new CFI formulas for volume, basal area and DBH class. USDA Forest Service Region 9, Forest Control by Continuous Inventory Note No. 97. 6 pp.

Wenzel, L.C., J. Levitan, and K. Barber. 1980. Selection of basal area factor in point sampling. Journal of Forestry 78:83-84.

Zeide, B. 1980. Plot size optimization. Forest Science 26:251-257.

Appendix 1 - Sample sizes for double sampling with fixed cost

Oderwald and Jones (1992) present double sampling sample size formulas for the case of equal precision (of that of a simple random sample). Here we take the same approach to derive formulas for n_L and n_S that give minimum variance for double sampling with cost equal to a simple random sample. The initial formulas are all from Oderwald and Jones (1992).

The cost of implementing a simple random sample of size n_e is

$$c_S * n_e \qquad [A1-1]$$

where c_s is the cost of observing volume per acre at a sample location. The cost of implementing a double sample where trees are measured at n_S locations and trees are only tallied at $n_L - n_S$ locations is

$$c_S\left(\frac{n_L - n_S}{K} + n_S\right) \qquad [A1-2]$$

where K is the ratio of time to observe volume per acre to time to observe basal area per acre at a sample location.

The variance of the double sampling estimator, expressed relative to the mean, is

$$\frac{CV_G^2\left(B^2 - 2\rho B + 1\right)}{n_S} + \frac{CV_V^2 - CV_G^2\left(B^2 - 2\rho B + 1\right)}{n_L} \qquad [A1-3]$$

where

CV_G = coefficient of variation of basal are per acre
CV_V = coefficient of variation of volume per acre
ρ = correlation coefficient between volume per acre and basal area per acre
B = ratio of the coefficient of variation of volume per acre to the coefficient of variation of basal area per acre

The solution sought minimizes [A1-3] subject to [A1-2] equal to [A1-1]. Applying the Lagrange multiplier method gives

$$\begin{aligned} n_L &= n_e K \frac{\sqrt{W}}{1+\sqrt{W}} \\ n_S &= n_e K \frac{1}{(K-1)\left(1+\sqrt{W}\right)} \end{aligned} \qquad [A1-4]$$

where

$$W = \frac{(2\rho B - 1)}{(B^2 - 2\rho B + 1)(K - 1)} \quad \text{[A1-5]}$$

The sample sizes [A1-4] can be substituted into [A1-3] to find the double sampling minimum variance. Expressing that result relative to the variance of the equal cost simple random sample gives the realized efficiency of the double sample. Alternatively, the two variances can be equated and the value of K at which they are equal can be solved for.

Note that

$$\frac{n_S}{n_L} = \sqrt{\frac{B^2 - 2\rho B + 1}{(2\rho B - 1)(k - 1)}} \quad \text{[A1-6]}$$

This is the same ratio that Oderwald and Jones (1992) reported. The fraction of plots on which trees are measured is the same for both the fixed cost and fixed precision solutions.

Appendix 2 - Sample sizes for and efficiency of the Big BAF method

The results provided here assume a level of mathematical sophistication significantly beyond the remainder of the presentation. They are provided for completeness as it appears they have not previously been presented in the literature. The initial derivations follow the approach of Oderwald and Jones (1992) for double sampling with a ratio-of-means estimator and variable-radius plots.

Optimal number of plots and number of trees (fixed precision)

Our estimator of volume per acre is (duplicate of equation [3-1])

$$V_{BBAF} = \overline{VBAR}_{BBAF} \cdot \overline{G}_{BBAF} \qquad [A2\text{-}1]$$

where $\overline{VBAR}_{BBAF}$ is the per tree average VBAR for trees selected with BAF_B and $\overline{G}_{BBAF}$ is the standard basal area per acre estimator from the plots tallied with BAF_S. We assume that BAF_B is chosen to be large enough that the VBARs can be assumed i.i.d. and that $\overline{VBAR}_{BBAF}$ and $\overline{G}_{BBAF}$ are independent.

Goodman (1960) presented an exact formula for the true variance of the product of two independent estimators. Applying that result to [A2-1] gives

$$\frac{Var(V_{BBAF})}{V_{BBAF}^2} = \frac{Var(\overline{VBAR})}{\overline{VBAR}^2} + \frac{Var(\overline{G})}{\overline{G}^2} + \frac{Var(\overline{VBAR}) \cdot Var(\overline{G})}{\overline{VBAR}^2 \cdot \overline{G}^2} \qquad [A2\text{-}2]$$

where Var(*) means variance of * and we have dropped the BBAF subscript identifiers on $\overline{VBAR}_{BBAF}$ and $\overline{G}_{BBAF}$ to conserve space. [A2-2] can be rewritten in terms of coefficients of variation as

$$\frac{Var(V_{BBAF})}{V_{BBAF}^2} = \frac{CV_{VBAR}^2}{T_{BBAF}} + \frac{CV_G^2}{n_{BBAF}} + \frac{CV_{VBAR}^2 \cdot CV_G^2}{T_{BBAF} \cdot n_{BBAF}} \qquad [A2\text{-}3]$$

where T_{BBAF} is the total number of trees measured and n_{BBAF} is the number of plots, both for the Big BAF method, and CV_{VBAR} and CV_G are the coefficients of variation of the per tree VBARs and per plot Gs, respectively. To obtain the same precision as a simple random sample of size n_e when volume per acre coefficient of variation is CV_V we need (cf. 1-11]

$$\frac{CV_V^2}{n_e} = \frac{CV_{VBAR}^2}{T_{BBAF}} + \frac{CV_G^2}{n_{BBAF}} + \frac{CV_{VBAR}^2 \cdot CV_G^2}{T_{BBAF} \cdot n_{BBAF}} \qquad [A2\text{-}4]$$

or

$$\frac{B^2}{n_e} = \frac{U^2}{T_{BBAF}} + \frac{1}{n_{BBAF}} + \frac{CV_{VBAR}^2}{T_{BBAF} \cdot n_{BBAF}} \qquad [A2\text{-}5]$$

where

$$B = \frac{CV_V}{CV_G}$$
$$U = \frac{CV_{VBAR}}{CV_G} \qquad [A2-6]$$

and it is assumed that the BAF used in the simple random sample is identical to BAF_S.

The cost of a Big BAF sample is

$$COST_{BBAF} = c_S n_{BBAF} + c_B n_{BBAF} + c_m T_{BBAF} \qquad [A2-7]$$

where

c_S = cost of tallying a plot with BAF_S

c_B = cost of tallying a plot with BAF_B

c_m = cost of measuring a tree once it is determined "in"

The desired minimum cost solution for Big BAF (optimal T_{BBAF} and n_{BBAF}) can be found by minimizing [A2-7] subject to the constraint [A2-5] using the Lagrange multiplier method. The solution is

$$T_{BBAF}^{opt} = \frac{\sqrt{\frac{c_S + c_B}{c_m}} \sqrt{n_e\left(U^2 n_e + B^2 CV_{VBAR}^2\right)} + n_e U^2}{B^2}$$
$$n_{BBAF}^{opt} = \frac{n_e\left(T_{BBAF}^{opt} + CV_{VBAR}^2\right)}{\left(T_{BBAF}^{opt} B^2 - n_e U^2\right)} \qquad [A2-8]$$

The CV_{VBAR}^2 term in [A2-8] makes planning Big BAF efforts difficult. Experience shows that the term will seldom exceed 1.0 in value. Careful examination of [A2-8] indicates the term has minimal impact on the solution. Given the imprecision of our knowledge of the terms in [A2-8], from a practical planning perspective it is recommended one assume $CV_{VBAR}^2 = 1$ in [A2-8].

Re-expressing costs

The cost ratio in [A2-8] is not particularly straightforward to interpret or apply. The constituent costs (times) can be approximated based on the following relations

$$c_B = c_S \left(\frac{BAF_S}{BAF_B}\right)^{3/4} \qquad [A2-9]$$

$$c_m = \frac{k \cdot c_B BAF_B}{\tilde{G}} \qquad [A2-10]$$

where

$$k = \frac{\text{cost of measuring trees identified as "in" at a location}}{\text{cost of tallying trees at a location}}$$

$\tilde{G}$ = expected average basal area per acre

Equation [A2-9] is based on the analysis in Zeide (1980) and [A2-10] uses an expected number of measure trees per plot. The cost ratio then becomes

$$\frac{c_S + c_B}{c_m} = \frac{\tilde{G}}{k \cdot BAF_B}\left[\left(\frac{BAF_B}{BAF_S}\right)^{3/4} + 1\right] \qquad \text{[A2-11]}$$

Equation [A2-11] can be substituted into [A2-8]. However, the solution to [A2-8] implies a BAF_B

$$\frac{T_{BBAF}}{n_{BBAF}} \equiv \frac{\tilde{G}}{BAF_B} \qquad \text{[A2-12]}$$

The solution to [A2-8] thus involves iteration of the following steps

1. Specify B, U, BAF_S, k, $\tilde{G}$, and n_e
2. Guess BAF_B, solve [A2-8], and check [A2-12]
3. Refine BAF_B until [A2-12] is an equality

Cost comparison with simple random sampling

Substituting [A2-9] and [A2-10] into [A2-7] gives the minimum cost Big BAF solution

$$COST_{BBAF} = c_S \cdot \left(n_{BBAF}^{opt} + n_{BBAF}^{opt}\left(\frac{BAF_S}{BAF_B}\right)^{3/4} + \frac{k \cdot BAF_B \cdot T_{BBAF}^{opt}}{\tilde{G}}\left(\frac{BAF_S}{BAF_B}\right)^{3/4} \right) \qquad \text{[A2-13]}$$

The cost of a simple random sample of the same precision, implemented with BAF_S is

$$COST_{SRS} = n_e \cdot c_S \cdot (1 + k) \qquad \text{[A2-14]}$$

The ratio of [A2-13] to [A2-14] indicates the cost efficiency advantage of the Big BAF method. It is a function of B, U, the expected number of tally trees per plot with BAF_S, k, and n_e. The value of k for which $COST_{BBAF} = COST_{SRS}$ can be found by iteration with the previously described iteration steps imbedded; larger values of k imply a cost efficiency gain.

Optimal number of plots and number of trees (fixed cost)

In some instances, interest will lie in the Big BAF sample sizes that give minimum variance for cost equal to a simple random sample. The optimal equal cost solution is found by minimizing the right hand side of [A2-5] subject to the constraint that [A2-7] is equal to [A2-14]. Applying the Lagrange multiplier method results in the following quadratic equation

$$\mathrm{A}\left(T_{BBAF}^{opt}\right)^2 + \mathrm{B}T_{BBAF}^{opt} + \mathrm{X} = 0 \qquad \text{[A2-15]}$$

where

$$\mathrm{A} = \frac{c_m}{c_S + c_B} U^2 - 1$$

$$\mathrm{B} = -2\left(\frac{c_S}{c_S + c_B} n_e (1+k) U^2 + CV_{VBAR}^2 \right)$$

$$\mathrm{X} = \frac{c_S}{c_m} \frac{c_S}{c_S + c_B} n_e^2 (1+k)^2 U^2 + \frac{c_S}{c_m} n_e (1+k) CV_{VBAR}^2$$

Once T_{BBAF}^{opt} is found using [A2-15] n_{BBAF}^{opt} can be found using

$$n_{BBAF}^{opt} = \frac{c_S}{c_S + c_B} n_e (1+k) - T_{BBAF}^{opt} \frac{c_m}{c_s + c_B} \qquad \text{[A2-16]}$$

Solution of [A2-15] and [A2-16] requires the same re-expression of costs and iteration as the fixed precision approach. CV_{VBAR}^2 can again safely be assumed to be 1.0. Note, significantly though, that the solution does not involve B. The value of B will only determine the statistical efficiency of Big BAF over SRS (see below) for the equal costs solution.

Substituting T_{BBAF}^{opt} and n_{BBAF}^{opt} into the right-hand-side of [A2-5] gives the minimum Big BAF sample (relative) variance which can be compared to the left-hand-side of [A2-5] to determine the efficiency of the optimal Big BAF sampling solution. Or, the two (relative) variances can be equated to obtain the value of k for which the two precisions are equal.