Location Uncertainty and the Tri-Areal Design

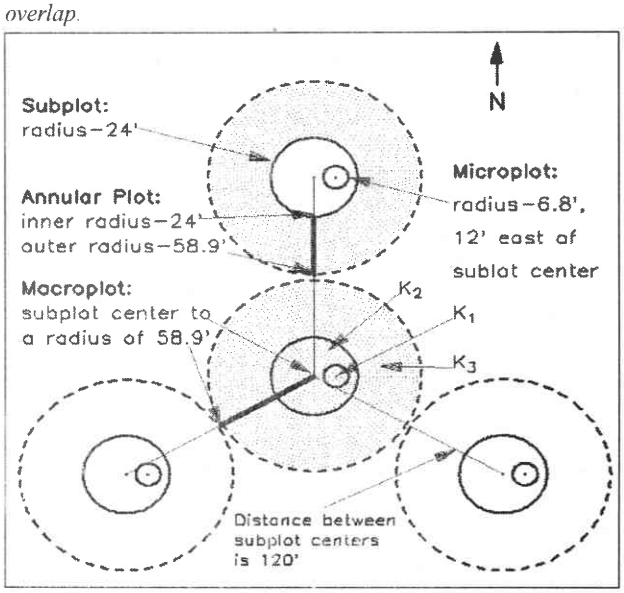
Francis A. Roesch[1]

Abstract.—The U.S. Department of Agriculture Forest Service Forest Inventory and Analysis Program (FIA) uses a field plot design that incorporates multiple sample selection mechanisms. Not all of the five FIA units currently use the entire suite of available sample selection mechanisms. These sampling selection mechanisms could be described in a number of ways with respect to the optional mechanism known as the annular plot. The annular plot is an auxiliary sampling mechanism intended for sampling rare attributes of interest. One explanation is that the subplot, which samples all trees greater than or equal to 5 in diameter at breast height (d.b.h.), is surrounded by an annular plot, concentric with the subplot for the estimation of rare but regionally important events. To date this selection mechanism has only been used to increase the sample of larger trees above a predefined d.b.h., known as a breakpoint diameter. Alternatively, the selection mechanisms could be viewed as disjoint concentric circles. The subplot in this latter view would sample all trees that are greater than or equal to 5 in and less than the breakpoint diameter. The larger circle can be referred to as a macroplot and it serves as the sole sampling mechanism for trees greater than or equal to the breakpoint diameter. This article focuses on the importance of clarity between these two descriptions and the estimation bias that can result from a misunderstanding of the distinctions between them, especially with respect to change estimates.

Introduction

The U.S. Department of Agriculture Forest Service Forest Inventory and Analysis (FIA) program uses a field plot design that is fairly represented by figure 1. The sampling selection mechanisms represented by the plot design could be described in a number of ways with respect to the annular plot portion of the design. The annular plot is an auxiliary sampling mechanism for rare attributes of interest. One explanation is that the subplot samples all tree greater than or equal to 5 in diameter at breast height (d.b.h.) and is enclosed by a circle of radius 24 ft. The annular plot is concentric with the subplot, beginning at a distance of 24 ft from subplot center and ending at 58.9 ft from subplot center, forming an annulus around the subplot. FIA allows this selection mechanism for rare, but regionally important, events. Until now, the annular plot has only been used to increase the sample of larger trees with a

Figure 1.—*The Forest Inventory and Analysis plot design, showing the annular plot view, in which the sample areas are disjoint, and the macroplot view, which is analogous to discrete horizontal point sampling. In the later view, the sample areas overlap.*



[1] Mathematical Statistician, U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis, Southern Research Station, Asheville, NC.

predefined d.b.h., known as a breakpoint diameter. The annular plot view treats the microplot and subplot as the primary sample and the annular plot, consisting of an annulus around the subplot, as an auxiliary sample.

Alternatively, the selection mechanisms could be viewed as disjoint overlapping circles. The subplot in this case would sample all trees that are greater than or equal to 5 in and less than the breakpoint diameter. The larger circle can be referred to as a macroplot, and it constitutes the entire sample of trees greater than or equal to the breakpoint diameter. For simplicity we'll call the former description the annular plot view and the latter the macroplot view. This article focuses on the importance of clarity between these two descriptions and the estimation bias that can result from a misunderstanding of the distinctions between them.

Bechtold and Patterson (2005) define both the bi-areal and the tri-areal plot designs. In relation to the development found there, the annular plot view would hold that the bi-areal design is common throughout the United States and that some regions may choose to include an auxiliary sample collected on annuli surrounding each of the subplots. The macroplot view differs in that each region applies either a bi-areal or a tri-areal design, which coincide exactly only by the definition of the sample selected from the microplot. That is, the subplot samples a different population partition in the tri-areal design than in the bi-areal design, with the macroplot existing in only the tri-areal design. The estimators for a single point in time given in Bechtold and Patterson (2005) can be derived through either view. A practical advantage of the annular plot view is that if the auxiliary sample is not conducted in some regions, then the entire population is still sampled by an identical primary sample.

Enter the Temporal Dimension

An important class of variables exists for which a partitioning of the macroplot into the inner macroplot (equal to the area of the subplot) and the outer macroplot (equal to the area of the annular plot) is necessary. That class of variables consists of those whose ranges are to be partitioned by the various selection mechanisms, and whose measures can change over time. D.b.h. is possibly the most important member of this class of variables.

It is easy to show that from an instantaneous point of view, the macroplot view and the annular plot view both define a probability sample. FIA, however, measures a temporally continuous rather than an instantaneous population. From this perspective, as stated previously, we see that when the macroplot is used to sample large trees (say those with d.b.h. ≥ 25 in), it constitutes three distinct samples of that population. These three samples are (1) a sample (K_1) of trees that have been measured since they attained the 1-in class, selected with probability k_1 (proportional to the area of the microplot); (2) a sample (K_2) of trees that have been measured since they attained the 5-in class, selected with probability k_2 (proportional to the area of the subplot minus the area of the microplot); and (3) a sample (K_3) of trees that have been measured since they attained the breakpoint diameter, selected with probability k_3 (proportional to the area of the annular plot). Ignoring or not explicitly acknowledging the distinction between these samples has the potential to bias estimators of survivor value growth by two of the three primary compatible estimation systems published for remeasured horizontal point samples, due to the resulting location uncertainty. I'll follow others and refer to these three systems as Beers-Miller (Beers and Miller 1964), Van Deusen (Van Deusen *et al.* 1986), and Roesch (Roesch 1988, 1990; Roesch *et al.* 1989, 1991, 1993) estimators. The problem arises implicitly rather than explicitly in the first two systems, because, if data are collected and stored under the macroplot view, and there is a non-zero quality-control tolerance for horizontal distance from plot center, there would be no way of determining for certain whether a large tree recorded as physically located near the edge of the subplot and previously unrecorded was previously missed, previously smaller than 5 in d.b.h., or actually on the annular plot. This knowledge is necessary for strict application of the Beers-Miller and Van Deusen survivor growth estimators because they both require use of time 1 inclusion probabilities. The Roesch survivor growth estimator relies on time 2 inclusion probabilities and therefore would sidestep the problem.

The survivor sample ("s") consists of sample trees measured and above a merchantability limit on consecutive occasions, while the new sample ("n") consists of trees that were above the merchantability limit on both occasions, but eligible to be sampled for the first time on the second occasion. For clarity we'll define the estimators:

Beers-Miller: $\hat{S}_B = s_2' - s_1$

Van Deusen: $\hat{S}_V = s_2 - s_1 + n_2$

Roesch: $\hat{S}_R = s_2 - s_1' + n_2 - n_1'$

where:

S_1 = estimate of time 1 value of the "s" sample using time 1 inclusion probabilities.

S_1' = estimate of time 1 value of the "s" sample using time 2 inclusion probabilities.

S_2 = estimate of time 2 value of the "s" sample using time 2 inclusion probabilities.

S_2' = estimate of time 2 value of the "s" sample using time 1 inclusion probabilities.

n_2 = estimate of time 2 value of the "n" sample using time 2 inclusion probabilities.

n_1' = model estimate of time 1 value of the "n" sample with time 2 inclusion probabilities.

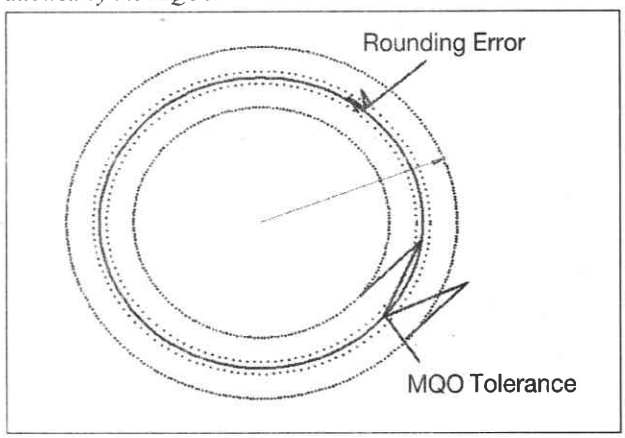
As a point-sampling estimator of survivor growth, $\hat{S}_R$ was shown in the citations above through simulations to dominate the other estimators in terms of squared error loss. It does, however, require predictions of time 1 values that are not required of the other two estimators in the absence of location error. In the absence of location error, all three estimators are unbiased estimators for survivor growth (Beers and Miller 1964, Van Deusen *et al.* 1986, Roesch 1988). We will show below that location error contributes to a bias in the Beers-Miller estimator and the Van Deusen estimator (but not in the Roesch estimator) through two mechanisms:

(1) Rounding error.

(2) Measurement error allowed by the measurement quality objective (MQO).

These errors are depicted in figure 2. Rounding error contributes positive bias to the inclusion probability, while measurement error allowed by the MQO contributes bias and variance. The bias from measurement error is due simply to a symmetric linear error being applied to a point on the radius of a circle.

Figure 2.—*Two potential errors associated with measurement along a radius: rounding to a discrete result and the tolerance allowed by the MQOs.*



MQO = measurement quality objective.

No Measurement Error

Assume that the field crew is not required to determine of the location of trees with respect to K_1, K_2, and K_3, leaving sample assignment to be inferred from the distance from subplot or microplot center, which is rounded and recorded to the nearest $1/t$ foot, where t is a positive integer. This use of rounded distances will result in sample trees that appear to be in the "s" sample that are actually in the "n" sample, thus creating "apparently missed" trees.

For now, also assume that our concern is with merchantable value growth of above-threshold trees. In this case we can ignore the microplot and concentrate on the effects of the tolerance definitions at the border between the subplot and the annular plot. Distance measures are continuous variables that are recorded in discrete units. A recording of distance d would result from rounding of the true distance D in the interval $\left(d - \frac{1}{2t}\right) < D \leq \left(d + \frac{1}{2t}\right)$ feet. The difference between D and d is known as rounding error. Assume trees are randomly distributed over the land area. Then D is randomly distributed within the

annulus bounded on the inside by $\left(d-\frac{1}{2t}\right)$ and on the outside by $\left(d+\frac{1}{2t}\right)$. This annulus ($A$) has an area of:

$$\pi\left[\left(d+\frac{1}{2t}\right)^2-\left(d-\frac{1}{2t}\right)^2\right]=\frac{2\pi d}{t}. \qquad (1)$$

D is distributed within the annulus as:
$p(D) \sim 1/(2\pi D)$.

Example

Set $d = 24$ ft. and $t = 10$, then the annulus has an area of $\frac{2\pi d}{t}=\frac{2\pi 24}{10}=4.8\pi \text{ft}^2$.

This would appear small relative to the nominal area of the subplot, which is $\pi r^2=\pi 24^2 \approx 1809.56\text{ft.}^2$ The area outside of the subplot boundary that would appear inside of the boundary after rounding lies between 24 and 24.05 ft and is equal to:

$$\pi\left[\left(24+\frac{1}{2t}\right)^2-(24)^2\right]\text{ft}^2=\pi\left[\frac{24}{t}+\frac{1}{4t^2}\right]\text{ft}^2=$$
$$\pi\left[2.4+0.0025\right]\text{ft}^2=2.4025\pi\,\text{ft}^2$$

resulting in an unrecognized selection bias due to rounding error of $b_S^r=\frac{2.4025}{576}\approx 0.417$ percent.

In a strict application of the Beers-Miller estimator ($\hat{S}_B=s'_2-s_1$) in the presence of rounding error, we might assume the apparently missed trees (truly members of the "n" sample) were actually missed (i.e., apparently members of the "s" sample), and subtract an estimate of time 1 value from their time 2 value, expand that by the inverse of the subplot area, and add it to the estimate of survivor growth, creating a positive bias (b_B^r), because the value growth of these trees should not have been included in the survivor growth estimate. Therefore the expected value of the Beers-Miller estimator in the presence of rounding error is:

$$E\left(\hat{S}_B^r\right)=E\left(\hat{S}_B+b_B^r\right)=E\left(s'_2-s_1\right)+E\left(b_B^r\right)\approx 1.00417\left(s'_2-s_1\right)$$

because the expected value of the bias due to rounding error is $E\left(b_B^r\right)\approx 0.00417\left(s'_2-s_1\right)$.

Conversely, in a strict application of the Van Deusen estimator ($\hat{S}_V=s_2+n_2-s_1$), we would mistakenly subtract an estimate of time 1 value expanded by the inverse of the subplot from the time 2 value estimate expanded by the inverse of the macroplot. The result will often be a relatively large negative number. Because the apparently-missed tree was actually in the "n" sample, no time 1 value should have been subtracted, therefore a negative bias results. The expected value of the Van Deusen estimator in the presence of rounding error in this case is:

$$E\left(\hat{S}_V^r\right)=E\left(\hat{S}_V+b_V^r\right)=E\left(s_2-s_1+n_2\right)+E\left(b_V^r\right)$$
$$=E\left(s_2-s_1+n_2\right)+\left(E\left(b_B^r\right)-0.00417E\left(s'_2\right)\right)$$

because the expected value of the bias is:

$$E\left(b_V^r\right)=\frac{\pi\left(24.05^2-24^2\right)}{\pi\left(58.9^2-24^2\right)}E\left(-n_1\right)$$
$$=\left[\frac{\pi\left(24.05^2-24^2\right)}{\pi\left(58.9^2-24^2\right)}\right]\left[-\frac{\pi\left(58.9^2-24^2\right)}{\pi 24^2}E\left(s_1\right)\right]$$
$$\approx -0.00417\left(s_1\right)$$

To use the Roesch estimator ($\hat{S}_R=s_2-s'_1+n_2-n'_1$), in the presence of rounding error, the apparently missed trees would be treated in the same manner as the trees in the "n" sample even though they appear to belong in the "s" sample. That is, we would predict the time 1 value and subtract it from the time 2 value, and then multiply the result by the inverse of the time 2 inclusion probability, so the selection bias would not affect the survivor growth estimator.

In the Presence of Measurement Error

Suppose that the following quality control standards for horizontal distance are enforced:

Tolerance:

Microplot: $\pm\ t_1$ ft

Subplot: $\pm\ t_2$ ft from 0-22.9 feet; $\pm\ t_3$ foot for > 23 feet

Macroplot: $\pm\ t_4$ ft

To simplify the discussion, we'll also assume that the tolerance is to be met 100 percent of the time and our concern is with merchantable value growth. Therefore, we can ignore the microplot and concentrate on the effects of the tolerance

definitions at the border between the subplot and the annular plot. A recording of distance d could result from a true distance D in the interval $\left(d - t_4 - \frac{1}{2t}\right) < D \leq \left(d + t_4 + \frac{1}{2t}\right)$ feet. Assume trees are randomly distributed over the land area.

This annulus (A) has an area of:

$$\pi\left[\left(d + t_4 + \frac{1}{2t}\right)^2 - \left(d - t_4 - \frac{1}{2t}\right)^2\right] = 2\pi d\left(2t_4 + \frac{1}{t}\right). \qquad (2)$$

Again D is distributed within the annulus:
$p(D) \sim 1/(2\pi D)$.

The additional error in (2), above rounding error in (1), is within-tolerance measurement error.

If we assume that the horizontal distance error follows any symmetric distribution, more trees will mistakenly appear to be inside of the division line than outside of the division line. Therefore, to calculate the allowed measurement error selection bias, note that we must first calculate the ratio of annular area external to the division line to the annular area internal to the division line:

$$R(d,t,t_4) = \frac{\pi\left[\left(d+t_4+\frac{1}{2t}\right)^2 - \left(d+\frac{1}{2t}\right)^2\right]}{\pi\left[\left(d-\frac{1}{2t}\right)^2 - \left(d-t_4-\frac{1}{2t}\right)^2\right]} = \frac{\left(2d+t_4+\frac{1}{t}\right)}{\left(2d-t_4-\frac{1}{t}\right)}$$

This ratio R is then applied to the selection area to determine the bias:

$$b(d,t,t_4) = \frac{\left[\frac{R(d,t,t_4)+1}{2}\right]\pi\left[\left(d-\frac{1}{2t}\right)^2 - \left(d-t_4-\frac{1}{2t}\right)^2\right] + \pi\left[\left(d^2 - \left(d-\frac{1}{2t}\right)^2\right) + \left(d-t_4-\frac{1}{2t}\right)^2\right]}{\pi d^2}$$

$$= \frac{\left[1-R(d,t,t_4)\right]\left[\frac{t_4^2}{2} + \frac{t_4}{2t} - dt_4\right]}{d^2}$$

Table 1 shows the selection bias (b) at 24 ft and its effect on the expected value of the bias in the Beers-Miller estimator when horizontal distance, rounded to 0.1 ft at five MQO tolerances of t_4 equal to 0.1, .05, 1.0, 3.0, and 6.0 ft, is used to determine tree location.

Table 1.—*The selection bias (b) at 24 ft and its effect on the expected value of the bias in the Beers-Miller estimator when horizontal distance is rounded to 0.1 ft at each MQO tolerance (t_4).*

Selection bias effect on the Beers-Miller estimator for FIA		
t_4	$b(d=24,t=10,t_4)$	$E(b_B^{rt})$
0.1	0.00002	$0.00419(s'_2-s_1)$
0.5	0.00042	$0.00459(s'_2-s_1)$
1.0	0.00173	$0.00590(s'_2-s_1)$
3.0	0.01555	$0.01973(s'_2-s_1)$
6.0	0.06220	$0.06637(s'_2-s_1)$

MQO = measurement quality objective.

Conclusions

We have noted that even a very accurate horizontal distance measurement will contribute an unnecessary bias of about 0.5 percent, because a distance of 24.05 ft would be recorded as 24 ft. The last two rows of table 1, where t_4 equals 3 and 6 ft, clearly show that horizontal distance tolerances that would be quite reasonable if one merely intended to be able to relocate very large trees would contribute to substantial bias if that same measure were later used to determine sample membership with respect to K_2 and K_3. Conversely, a field determination of "in subplot" or "in annular plot" will have an equal variance to the

use of horizontal distance for this purpose, but should not be biased. This observation is of concern because the assignment of the trees to areas (subplot or annulus) might not be recorded under the macroplot view of the sample design but is critical for Beers-Miller and Van Deusen growth estimates. In the former (chosen for use by FIA), growth below the threshold diameter for trees currently above the threshold diameter is measured exclusively on the subplot, making it necessary to distinguish between the area of the subplot and that of the annular plot. The alternative design-based Van Deusen estimator is more efficient because it uses the growth information from the annular plot that is ignored by the Beers-Miller estimator. Under the conditions investigated, it would suffer from an expected bias equal in magnitude to that of the Beers-Miller estimator, but of opposite sign. The model-based Roesch estimator would not incur bias due to location error, given an unbiased time 1 estimator for trees in the "n" sample. These predictions contribute a variance component to the Roesch estimator that is not present in the other estimators. That additional variance component will usually be smaller than the variance reduction achieved by the use of auxiliary information.

Although the subplot-microplot analogy to the annular plot-subplot issue is not trivial in all cases, we haven't discussed it here because it has less effect on value growth estimation because most value equations have a result of 0 below 5 in d.b.h., and those that do not have a very low result.

The Beers-Miller and Van Deusen estimators are inherently unbiased estimators. Therefore, the bias discussed in this paper would be an unnecessary artifact of collecting and storing data under the macroplot view if caution were not taken to prevent the confounding of the K_1, K_2, and K_3 samples.

Literature Cited

Bechtold, W.A.; Patterson, P.L., eds. 2005. The enhanced Forest Inventory and Analysis program—national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.

Beers, T.W.; Miller, C.I. 1964. Point sampling: research results, theory, and applications. Purdue University Research Bulletin 786. West Lafeyette, IN: Purdue University. 56 p. and insert.

Roesch, F.A., Jr. 1988. Compatible estimators for the components of growth and instantaneous volume calculated from remeasured horizontal point samples. New Brunswick, NJ: Rutgers University. Ph.D. dissertation.

Roesch, F.A., Jr. 1990. The variances of some compatible estimators for the components of growth from remeasured point samples. In: LaBau, V.J.; Cunia, T., eds. State-of-the-art methodology of forest inventory: a symposium proceedings. PNW-GTR-263. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station: 189-196.

Roesch, F.A., Jr.; Green, E.J.; Scott, C.T. 1989. New compatible estimators for survivor growth and ingrowth from remeasured horizontal point samples. Forest Science. 35(2): 281-293.

Roesch, F.A., Jr.; Green, E.J.; Scott, C.T. 1991. Compatible basal area and number of trees estimators from remeasured horizontal point samples. Forest Science. 37(1): 136-145.

Roesch, F.A., Jr.; Green, E.J.; Scott, C.T. 1993. A test of alternative estimators for volume at time 1 from remeasured horizontal point samples. Canadian Journal of Forest Research. 23(4): 598-607.

Van Deusen, P.C.; Dell, T.R.; Thomas, C.E. 1986. Volume growth estimation from permanent horizontal points. Forest Science. 32(2): 415-422.