New Compatible Estimators for Survivor Growth and *Ingrowth* from Remeasured Horizontal Point Samples

Francis A. Roesch, Jr.

Edwin J. Green

Charles T. Scott

ABSTRACT. Forest volume growth between two measurements is often decomposed into the components of survivor growth (S), ingrowth (I), mortality (M), and cut (C) (for example, Beers 1962 or Van Deusen et al. 1986). Net change between volumes at times 1 and 2 $(V_2 - V_1)$ is then represented by the equation

$$V_2 - V_1 = S + I - M - C.$$

Two new compatible pairs of estimators for S and Z in this equation are presented for use with the usual estimators of V_2 and V_1 from permanent horizontal point samples. In addition, a new estimator for $V_1 (V_1^{**})$ is presented, which takes advantage of the data from time 2. These estimators are evaluated along with previously existing estimators over a range of sample sizes. In all but the smallest of inventories, one of the new estimators of $S (S^{**})$ will provide better estimates of survivor growth in more cases than any previously developed estimator of S. S^{**} is shown to be compatible with V_1^{**} and the Purdue estimators of ingrowth, while empirical evidence is presented to support earlier contentions that the Purdue estimator is the best possible estimator of ingrowth from point samples. Recommendations are given based upon the presence or absence of a compatibility requirement and willingness to change the estimate of volume at time 1. FOR. SCI. 35(2):281–293.

ADDITIONAL KEY WORDS. Point sampling, forest growth, forest inventory, compatible estimators.

FOREST VOLUME GROWTH between two measurement periods, say times 1 and 2, is often decomposed into the components of survivor growth (S), ingrowth (I), mortality (M), and cut (C) (for example, Beers 1962 or Van Deusen et al. 1986). Net change is then represented by the equation

$$V_2 - V_1 = S + I - M - C \qquad (1)$$

where V_i equals the volume at time i.

Martin (1982) defined six categories of trees encountered in remeasured point samples. Four of these categories included trees that were "in" the point samples at time 1. *Ingrowth* trees are those live trees that were below

minimum dbh (nonmerchantable) and "in" (sampled) at time 1 but are above minimum dbh (merchantable) by time 2. *Survivor* trees are those live trees "in" and above minimum dbh at both measurements. *Mortality* trees are those trees that were "in" and above minimum dbh at the first measurement but die before the second measurement. *Cut* trees are trees harvested between measurements but otherwise analogous to mortality trees.

The two final categories consist of trees that were "out" at the first measurement and alive and "in" at the second measurement. *Ongrowth* trees were below minimum dbh and "out" at the first measurement but above minimum dbh and "in" at the second. *Nongrowth* trees were above minimum dbh and "out" at the first measurement but grew suffkiently to be "in" at the second measurement.

Van Deusen et al. (1986), using methodology developed by Palley and Horwitz (1961), derived the probability of a tree's inclusion in a specific category given its merchantability status at time 1 and that the tree is alive at time 2. They then showed that the published estimators for survivor growth (including the one that they present) and ingrowth differed by the placement of an error term. This paper proposes new estimators for S, Z and V,. One combination of estimators is shown to subdivide that error term and assign each subdivision to the proposed "proper" component of the growth and volume estimators. The notation of Martin (1982) will be augmented, where necessary, to facilitate our discussion of the new estimators.

Martin defined:

i = value obtained from the final volumes and final basal areas of ingrowth trees;

i' = value obtained from the final volumes and initial basal areas of ingrowth trees;

o = value obtained from the final volumes and final basal areas of ingrowth trees;

n = value obtained from the final volumes and final basal areas of nongrowth trees;

s_1 = value obtained from the initial volumes and initial basal areas of survivor trees;

s_2 = value obtained from the final volumes and final basal areas of survivor trees;

s_2' = value obtained from the final volumes and initial basal areas of the survivor trees;

m = value obtained from the initial volumes and initial basal areas of mortality trees;

c = value obtained from the initial volumes and initial basal areas of cut trees.

We define the following:

n_2 = value obtained from the final volumes and final basal areas of nongrowth trees (Martin's n);

n_1 = value obtained from the initial volumes and initial basal areas of the nongrowth trees;

n_1' = value obtained from the initial volumes and final basal areas of nongrowth trees;

$s_1' =$ value obtained from the initial volumes and final basal areas of survivor trees.

For the moment we will use the usual estimators for the yield at times 1 and 2. These estimators, as presented by Van Deusen et al. (1986) but modified to reflect our augmented notation, are

$$\hat{V}_1 = s_1 + m + c, \tag{2}$$

and

$$\hat{V}_2 = s_2 + n_2 + i + o. \tag{3}$$

The change in volume at each sample point was shown to be estimable by

$$\hat{V}_2 - \hat{V}_1 = s_2 + n_2 + i + o - s_1 - m - c. \tag{4}$$

If we estimate the change in volume with (4) then, ignoring mortality and cut, for two estimators of Z and S to be compatible they must sum to the first **five** terms in the right hand side of (4). Martin used the usual estimator for survivor growth which holds the old plot sizes constant and uses only the s trees:

$$\hat{S} = s_2' - s_1. \tag{5}$$

Martin then presented an **ingrowth** estimator that is compatible with $\hat{S}$,

$$\hat{I}_A = i + o + n_2 + s_2 - s_2'. \tag{6}$$

The volume of the nongrowth trees at time 2 is included in this estimator of ingrowth. Van Deusen et al. (1986) recognized that, by definition, nongrowth trees are not ingrowth. They correctly grouped Martin's n trees into the survivor category when they developed the following unbiased estimator for S:

$$\tilde{S} = s_2 - s_1 + n_2. \tag{7}$$

They then showed that their new estimator for survivor growth was compatible with the "revised Purdue" estimator for **ingrowth** (Beers and Miller 1964):

$$\hat{I}_P = i + 0. \tag{8}$$

Note that in order to achieve unbiasedness all of the volume of the n trees at time 2 is used in $\tilde{S}$, instead of just using the growth of these trees from time 1 to time 2. Obviously, the growth of these trees is not known because the trees were not measured initially. This is a problem inherent in point sampling. If the volumes of these nongrowth trees could be reasonably estimated for time 1, then it would be possible to use only the growth of the sampled survivor trees *(n* and *s* trees) to estimate survivor growth (i.e., we would recognize and subtract the nongrowth portion of nongrowth). We propose two such estimators, one biased and one unbiased, and their compatible **ingrowth** counterparts. The expected values of the new estimators can be derived using the procedure shown by Van Deusen et al. (1986) and are available from the authors upon request. The biased estimator is

$$S^* = s_2 - s_1 + n_2 - n_1, \tag{9}$$

and its compatible **ingrowth** estimator (also biased) is

$$I^* = i + o + n_1. \tag{10}$$

S^* has several potentially undesirable characteristics. As already mentioned, S^* is biased. Also, the time 1 volumes and basal areas have to be estimated for the nongrowth trees. We therefore present an estimator that will be unbiased if the predictions of the time 1 volumes of the nongrowth trees are unbiased. Although we cannot eliminate the need to estimate time 1 volumes, estimated time 1 basal areas are only used to distinguish between ongrowth and nongrowth. The unbiased estimator is

$$S^{**} = s_2 - s_1' + n_2 - n_1', \tag{11}$$

which is compatible with

$$I^{**} = s_1' - s_1 + n_1' + i + o. \tag{12}$$

To achieve compatibility, the nongrowth portion of nongrowth (i.e., that portion already present at time 1) is now illogically included in the estimator of ingrowth, I^{**}.

Although it is not compatible with any of the survivor growth estimators, we will also examine the Furnival estimator of ingrowth (Martin 1982),

$$\hat{I}_F = i', \tag{13}$$

due to its simplicity.

Van Deusen et al. (1986) showed that the compatible pairs of estimators ($\hat{S}$ with I_A) and ($\tilde{S}$ with I_p) differed by the placement of an error term,

$$\epsilon = s_2 - s_2' + n_2, \tag{14}$$

which has an expected value of zero.

Van Deusen's proposal[1] to experiment with the placement of different weights on ϵ led to the realization that ϵ can be decomposed into two components, each with an expected value of zero:

$$\epsilon = \epsilon_1 + \epsilon_G. \tag{15}$$

In (15)

$$\epsilon_1 = \{s_1' - s_1 + n_1'\} \tag{16}$$

is the time 1 yield component and

$$\bullet \quad o = \{(s_2 - s_1') - (s_2' - s_1) + (n_2 - n_1')\} \tag{17}$$

is the growth component.

To verify (15), add the right hand sides (r.h.s.) of (16) and (17) and cancel to obtain the r.h.s. of (14). Since (15) is true, we only have to show that ϵ_1 has an expected value of zero to prove that both ϵ_1 and $\bullet$ G have an expected value of zero. The framework for this proof was provided by Van Deusen et al. [1986, Equations (9)–(11)]. Now

$$E\{\epsilon_1\} = E\{s_1' - s_1 + n_1'\}$$

$$= E\left[\frac{FA}{m} \sum_{j=1}^{m} \sum_{i \in M} \left[Y_{ij}\left[\frac{V_{1i}}{b_{2i}}\right] - Y_{ij}\left[\frac{V_{1i}}{b_{1i}}\right] + X_{ij}\left[\frac{V_{1i}}{b_{2i}}\right]\right]\right]$$

---

[1] Personal communication with Paul C. Van Deusen.

$$= FA \sum_{i \in M} \left[\frac{b_{1i}}{FA} \left[\frac{V_{1i}}{b_{2i}} \right] - \frac{b_{1i}}{FA} \left[\frac{V_{1i}}{b_{1i}} \right] + \frac{(b_{2i} - b_{1i})}{FA} \left[\frac{V_{1i}}{b_{2i}} \right] \right]$$

$$= 0,$$

where F is the basal area factor, A is the area of the forest, m is the number of sample points, and M is the set of trees alive at time 2 that were merchantable at time 1. Y_{ij} and X_{ij} are indicator variables for s tree i and n tree i, respectively, tallied from point j. V_{ki} and b_{ki} are the volume (or whatever tree characteristic is of interest) and basal area, respectively, of tree i at time k.

The preceding decomposition suggests that ϵ_1 should actually be included with the volume estimate at time 1 and not with the growth estimate. The estimate of V_1 should then be changed to the improved estimator.[2]

$$V_1{}^{**} = (\hat{V}_1 + \epsilon_1) = s_1' + n_1' + m + c, \tag{18}$$

and the estimate for change in volume becomes

$$\hat{V}_2 - V_1{}^{**} = s_2 - s_1 + i + o + n_2 - m - c - \epsilon_1$$

$$= \hat{I}_P + (\tilde{S} - \epsilon_1) - m - c$$

$$= \hat{I}_P + (\hat{S} + \bullet \quad G) - m - c$$

$$= \hat{I}_P + S^{**} - m - c. \tag{19}$$

In (19) S^{**} is compatible with $\hat{I}_P$. This compatibility is achieved, of course, by permitting a change in the estimate of volume at time one. This will occasionally conflict with the motive for desiring a compatible estimator in the first place.

EMPIRICAL COMPARISON OF THE ESTIMATORS

DATA DESCRIPTION

The data consisted of 58 fixed area, permanent plots established by the USDA Forest Service Northeastern Forest Experiment Station in Hancock County, Maine. Circular, concentric one-tenth acre pulpwood plots and one-fifth acre sawlog plots were established in 1968 and remeasured in 1981. All trees that were at least 5 in. in diameter at breast height (dbh) were measured on the inner tenth acre, while only trees of sawlog size (9 in. for softwoods and 11 in. for hardwoods) were measured on the outer tenth acre. Cubic foot and board foot volumes were calculated using individual tree volume equations developed by Scott (1979 and 1981). The fifth-acre plots were expanded to determine board-foot volume per acre on each plot, while the tenth-acre plots were expanded for cubic-foot volume per acre.

SAMPLING PROCEDURE

Point samples were drawn from the plot data, under the assumption that the expanded plot values were the true board foot and cubic foot volumes per acre for the area around each plot. We varied the Basal Area Factor (BAF), in increments of 5, from IO to 40 ft²/ac to determine the effect of relative plot

---

[2] Empirical evidence that $V,{}^{**}$ is an improved estimator of V, will be presented in a manuscript in preparation.

size on the behavior of the estimators (a BAF of 37.5 was also included because it is a standard factor used by the Forest Service).

The simulated point samples realistically represent the point sample that would have been taken from each plot center up to a distance of 52.7 ft (the radius of a fifth-acre plot). Not having sampled pulpwood-sized trees in the outer tenth acre had no effect on the point samples because, even at the smallest BAF of 10, a tree would have had to be larger than 13.5 in. dbh (i.e., substantially larger than minimum sawlog size) to be selected at a distance greater than 37.2 ft (the radius of a tenth-acre plot). However, the argument could be raised that there would have been trees farther than 52.7 ft which would have been in the point sample had it actually been taken in the field. Missing these trees might differentially bias the evaluation of the estimators of survivor growth and ingrowth. This problem was partially solved by assuming the forest surrounding the fifth acre has the same diameter distribution as that found on the fifth acre. This assumption allowed us to project our fifth-acre data out into the surrounding fifth acre.[3]

Since the trees in the sample were relatively small, the projection procedure only added 3 trees (2 s trees and $1n$ tree) at a BAF of 10, $1s$ tree at a BAF of 15, and no trees at the larger BAFs, over the 58 plots. There was no difference observed in the relative rankings of the estimators due to the projection procedure. This result supports the contention that drawing point samples from the fifth-acre plot data is a reasonably valid procedure.

ANALYSIS

We judged the estimators of ingrowth and survivor growth by how closely they estimated the plot values (the "truth") given (1) the presence or absence of additivity restrictions, and (2) the feasibility of changing the volume estimates made at time 1.

We examined the estimators using as much of the information contained on the permanent plots as possible. This approach eliminates some of the noise in the analysis, but it will also mask the effect of poor volume predictions on the growth estimators. The data description reveals that we knew the true board foot volume at time one for all of the point sample trees because all of the trees of sawlog size were measured, and the board foot volume of a tree smaller than sawlog size is zero. Obviously this is not the case for cubic foot volume and the cubic foot volumes of trees on the outer tenth acre which were not measured at time one had to be predicted (the actual method of prediction is available from the authors). Basal area had to be predicted for trees that were less than 5 in. dbh at time 1 on the inner tenth acre and those that were less than sawlog size at time 1 on the outer tenth acre. A simple linear regression of basal area at time 1 over basal area at time 2 of the s trees that were less than 0.75 ft^2 in basal area at time 2 was used. It was not the intent of this study to determine the best way to estimate the previous basal areas or volumes of the survivor trees in the sample that were not measured at time 1.

RESULTS AND DISCUSSION

Figure 1 includes two graphs, one for each type of volume, where $\bar{S}, \hat{I}_A, n_2$, $i, o,$ and $s_2 - s_1$ are summed over the 58 sample points and plotted against BAF. These graphs show that n_2 is the major component of $\bar{S}$ and that both

---

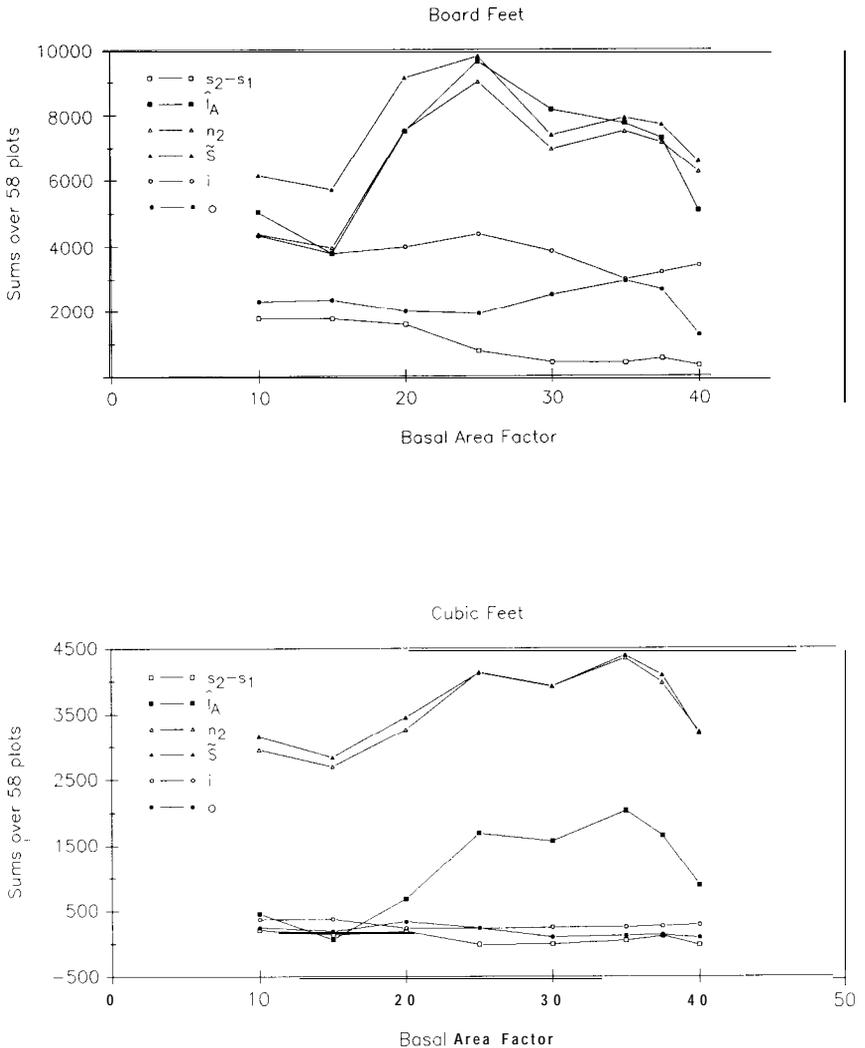[3] The method used to project the circular plots is available from the authors.

FIGURE 1. The sums, over 58 points, of $\bar{S}, \hat{I}_A$, and the components i, o, n,, and $s_2 - s_1$ versus BAF for Board Foot Volume (top) and Cubic Foot Volume (bottom).

$\bar{S}$ and $\hat{I}_A$ track n_2. The formulation of each of these two estimators allows the n trees, which were not measured for growth, to control the growth estimate.

The (1) squared differences, (2) absolute differences, and (3) arithmetic differences[4] between the estimate from each estimator and the "truth" on each plot were summed over the 58 plots for each BAF considered and will be referred to as the SSD, SAD, and SD, respectively. The SD indicates the bias of an estimator. The SAD and SSD are, in some sense, better measures of error because they include a bias component and a precision component. Figure 2 displays the three statistics plotted against BAF for each estimator of S, in terms of cubic feet. Figure 3 displays the same quantities where the unit of measurement is board foot. Similarly, Figure 4 shows the three sta-

---

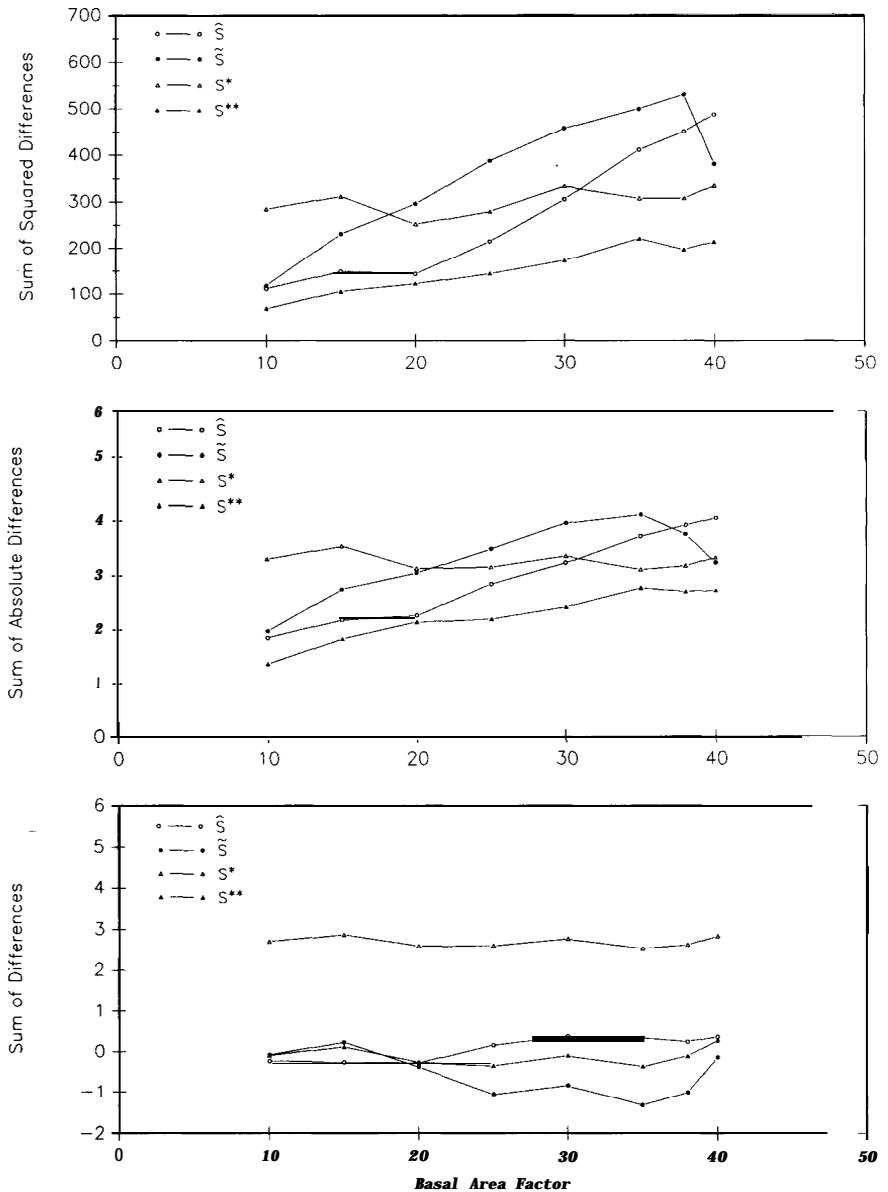[4] Difference = plot value − estimate, therefore a positive difference indicates underprediction.

FIGURE 2. The sums (x 10^{-3}) of squared differences, absolute differences, and differences between the estimators of cubic foot survivor growth and the fixed area plot values versus BAF.

tistics plotted against BAF for each estimator of I, where Z is measured in cubic feet, and Figure 5 is the board foot counterpart to Figure 4.

Figure 2 shows that, for cubic foot volume, S^{**} gives the best overall estimates of survivor growth. Both the sum of squared differences (SSD) and the sum of absolute differences (SAD) for S^{**} are the lowest over all of the BAFs studied. Also, the plot of the sum of differences (SD) shows that S^{**} has the least overall bias while $\hat{S}$ also displays a bias very close to zero. The bias of S appears a little less predictable, varying more widely around zero, while S^* exhibits a clear bias. However, S^* shows the least change in both the sum of absolute differences and sum of squared differences as BAF
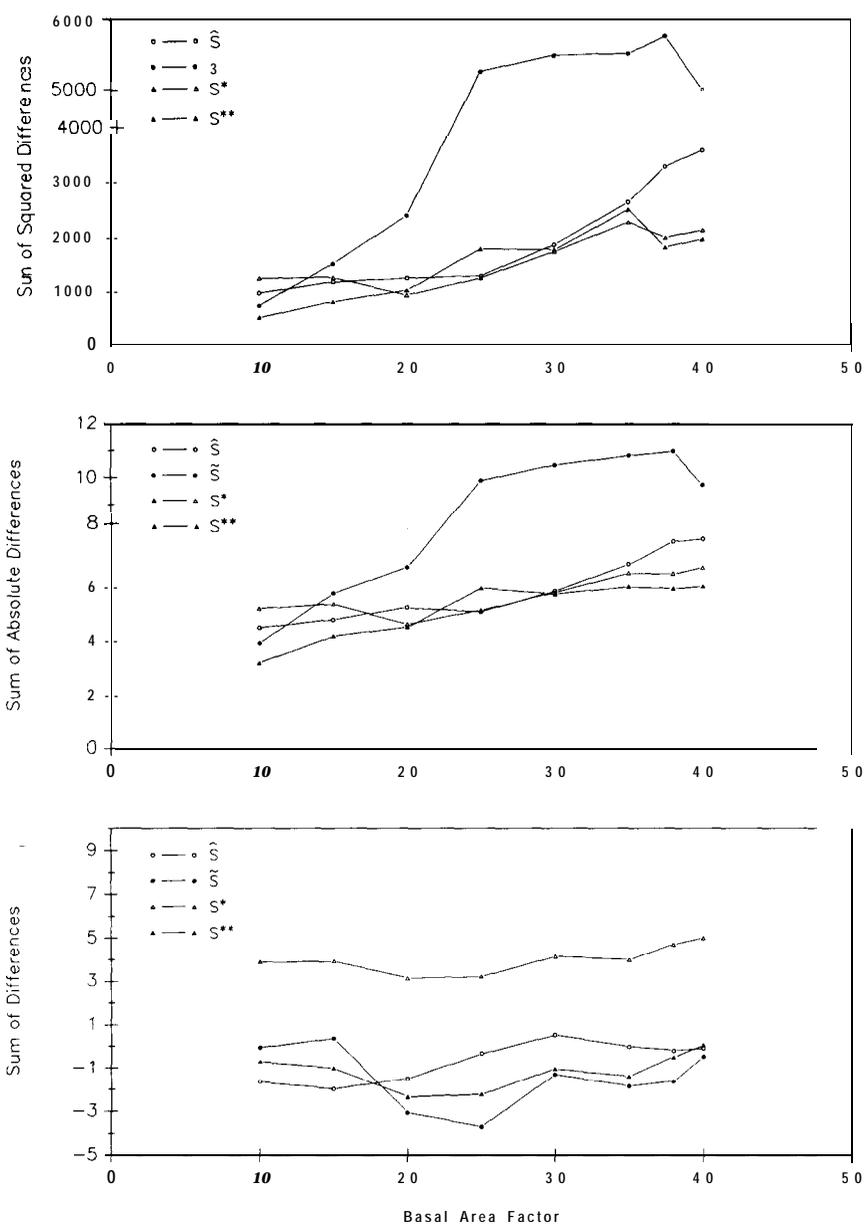
FIGURE 3. The sums ($\times 10^{-3}$) of squared differences, absolute differences, and differences between the estimators of board foot survivor growth and the fixed area plot values versus BAF.

increases. So the ranking of S* among the estimators improves as plot size increases.

The best estimator of board foot survivor growth is less clearcut than in the cubic foot case, although $\tilde{S}$ is clearly the worst at the larger BAFs. It can be seen in Figure 3 that at the extremes of plot size, S** has the lowest SSD and SAD, and S^{**}, S^*, and $\hat{S}$ exchange rankings by these two measures in the midrange of BAFs. The SD plot again shows the constant-bias of S^*, little difference between S** and $\hat{S}$, and a slightly more erratic S.

Of the ingrowth estimators, Figure 4 shows that none of the others are as good as $\hat{I}_F$ and $\hat{I}_p$. Of these two I_p is consistently slightly lower in both SSD and SAD. Since $\hat{I}_F$ is not compatible with any of the survivor growth estimators and would require additional measurement or estimation of time 1 basal areas (in the pulpwood case), there is no reason to choose it over I_p. The SSD and SAD plots also show that the ingrowth estimators which utilize nongrowth trees $(I^{**}, I^{*}, \text{ and } I_A)$ are ordered roughly by how much nongrowth they use. Z^{**} contains n_1', which places less weight on nongrowth than n_1 (because the basal area of a tree which grew is smaller at
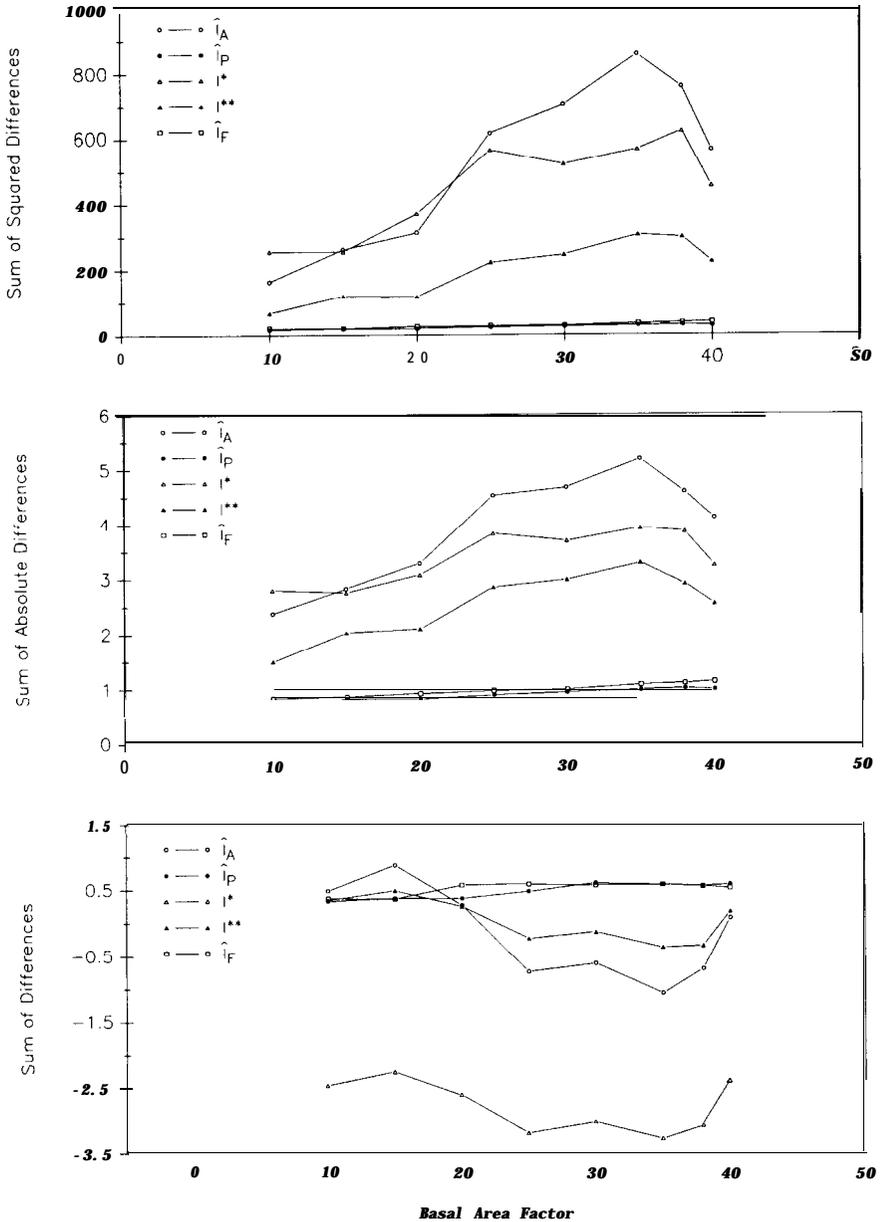

FIGURE 4. The sums ($\times 10^{-3}$) of squared differences, absolute differences, and differences between the estimators of cubic foot ingrowth and the fixed area plot values versus BAF.

time 1 than at time 2), and has the lowest overall SSD and SAD. I^* utilizes n_1 and has higher values for these measures, while $\hat{I}_A$ utilizes n_2 and has the highest SSD and SAD in most cases. Therefore, the more nongrowth in an ingrowth estimate the worse it becomes, as we should expect. Of the unbiased estimators, I^* has the least empirical bias in this analysis according to the SD plot. The bias for $\hat{I}_p$ and I_F is very small and predictable, while $\hat{I}_A$ varies more widely around zero than Z^{**} in much the same manner as $\hat{S}$ of the survivor growth estimates.

Examination of Figure 5 leads one to the same conclusions for board foot ingrowth as were drawn in the cubic foot case. The differences between $\hat{I}_p$ and $\hat{I}_A$ are more pronounced in the plots of SSD and SAD, while the differences between I^{**} and I_F are less pronounced. Again, the two ingrowth estimators that do not include nongrowth have a very slow, steady increase in both SSD and SAD as BAF goes up, while the other three estimators are much more erratic as they map the highly variable nongrowth estimates.

CONCLUSIONS

Of the two new compatible pairs of estimators for S and I in Equation (1), one of the new estimators (S^{**}) appears to provide better estimates of survivor growth in most cases than any previously developed estimator. Whether or not this is illusionary depends on how justified we were in using the known time 1 volumes of the n trees, when, in reality, they would have to be estimated. Our use of known time 1 volumes simplified this study immensely. Of the 17 species groups recognized by the USDA Forest Service Northeastern Forest Experiment Station as being sufficiently different to require individual volume equations, 15 were represented, in various proportions, in the data set. Because we varied plot sizes, the best method of predicting previous tree volumes for nongrowth trees would have varied not only by species group, but also by BAF. In inventories in which there are many more s trees than n trees of a particular size, a good relationship between time 1 and time 2 volumes should be easy to establish. Because this is the case for most inventories conducted in the Northeast, our simulation is probably realistic.

The recommendations that we will make assume that some reliable method of estimating time 1 volumes for the nongrowth trees can be found. The accuracy of these estimated volumes will decrease as the proportion of basal area at time 2 due to growth since time 1 increases, and if these volumes cannot be predicted with the desired degree of accuracy, then the (**) estimators should not be used at all. In large inventories, where there are a sufficient number of survivor trees of a particular species or species group, these predictions will most likely involve weighted regression of time 1 volumes on time 2 volumes. In small inventories or inventories in stands of great diversity, poorly represented species or species groups may have to be pooled or some additional information may have to be used, such as a previously developed yield equation.

Under the assumption above, three different sets of constraints on the estimators might be (1) additivity of the right-hand components to equal the left-hand components and the estimate of V_1 cannot be changed at time 2, (2) additivity is a requirement but the estimate for V_1 can be changed at time 2 and, (3) additivity is not a requirement.

The first set of constraints requires that all of the volume of nongrowth trees be included in the components of growth, even though not all of this volume is growth. To satisfy this set of constraints, one of the proposed
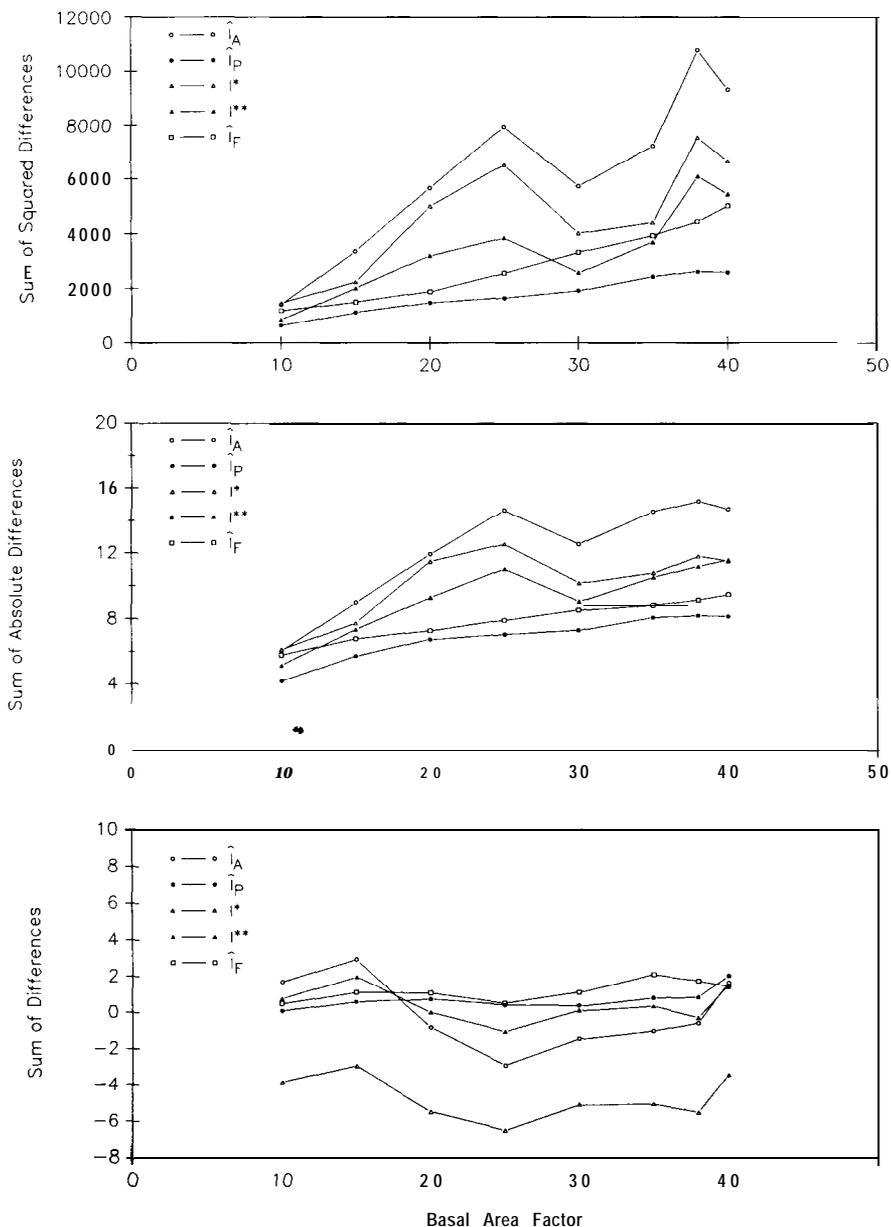
FIGURE 5. The sums ($\times 10^{-3}$) of squared differences, absolute differences, and differences between the estimators of board foot ingrowth and the fixed area plot values versus BAE

pairs of estimators (S^{**} with Z^{**}) is a reasonable alternative to the compatible pairs developed previously. Neither component should suffer from an extremely high variance and the survivor growth estimator (S^{**}) performed the best of those examined. Also, of the ingrowth estimators that contain nongrowth, Z^{**} was the most well behaved. If a higher priority is placed on the accuracy of the ingrowth estimator, then $\hat{I}_p$ should be used with $\tilde{S}$, although we have shown that $\tilde{S}$ could become unreliable due to its utilization of n_2. $\tilde{S}$, on the other hand, does not require the estimation of the volume of the nongrowth trees at time 1, as does S^{**}, and should improve relative to

S^{**} as the proportion of basal area at time 2 due to growth since time 1 increases.

The second set of constraints were that additivity is required but that V, could be changed. In this case, ϵ_1 should be included with the volume estimate at time 1 instead of one of the growth estimates. If this is done, the compatibility relationships change, and S^{**} becomes compatible with I_p. S^{**} could then be used with the Purdue estimator of ingrowth for the best estimate of each component.

Constraint (3) is the most liberal in that it allows the use of the best estimators for all of the parameters of interest by not requiring compatibility. However, the same conclusions are reached as for set (2).

The figures show that as more of Martin's n is included in an estimator of ingrowth, the variance of that estimator increases, and that as some threshold of this n is exceeded in an estimator of survivor growth, the variance of that estimator increases. We have suggested that part (in some cases a very large part) of this n, and therefore a part of the error due to n, doesn't belong to any type of growth, but rather to the volume present at time 1. The weakness of S^{**} and S^* is that basal areas and volumes of the nongrowth trees were not measured at time 1 and must therefore be estimated. Although the effect of the variance of these estimates on the growth and yield estimators should be small relative to the effect of ignoring the nongrowth portion of nongrowth, these effects should be investigated in general and for specific applications.

LITERATURE CITED

BEERS, T. W. 1962. Components of forest growth. J. For. 60(4):245–248.

BEERS, T. W., and C. 1. MILLER. 1964. Point sampling: Research results, theory, and applications. Purdue University Res. Bull. 786. 56 p. and insert.

MARTIN, G. L. 1982. A method for estimating ingrowth on permanent horizontal sample points. For. Sci. 28(1):110–114.

PALLEY, M. N., and L. G. HORWITZ. 1961. Properties of some random and systematic point sampling estimators. For. Sci. 7(1):52–65.

SCOTT, C. T. 1981. Northeastern forest survey revised cubic-foot volume equations. USDA For. Serv. Res. Note NE-304. 3 p.

SCOTT, C. T. 1979. Northeastern forest survey board-foot volume equations. USDA For. Serv. Res. Note NE-271. 3 p.

VAN DEUSEN, P. C., T. R. DELL, and C. E. THOMAS. 1986. Volume growth estimation from permanent horizontal points. For. Sci. 32:415–422.