THE FOURTH DIMENSION IN FIA

Francis A. Roesch[1]

Abstract.—In the past, the goal of forest inventory was to determine the extent of the timber resource. Predictions of how the resource was changing were made by comparing differences between successive inventories. The general view of the associated sample design included selection probabilities based on land area observed at a discrete point in time. That is, time was not considered part of the sample design because it was not considered an element of the sampled population. Over the last few decades, the general goal of Forest Inventory and Analysis (FIA) has been changing to monitoring the dynamic forest ecosystem. However, much of the literature discussing FIA's new annual monitoring system, its sample design, and estimators is still based on an areal probability paradigm. In Roesch (2008; Forest Science 54(4): 455-464), I pointed out why it is usually necessary to include the dimension of time when describing the sampled population and the sample design for FIA and similar forest inventory systems. Here, I further explore the inferential advantages of replacing the areal probability paradigm with a three-dimensional probability paradigm with an application.

INTRODUCTION

In the past, the primary goal of most national-scale forest inventories had been simply to determine the extent of the timber resource. Estimates of how the resource had changed, made by comparing differences between successive inventories, were merely an incidental benefit of these historic inventories. Over the last few decades, many national-scale forest inventories have morphed into full-fledged efforts to monitor many aspects of dynamic forest ecosystems. The U.S. Forest Service's Forest Inventory and Analysis Program (FIA) is no exception and is concerned with evaluating the dynamic state of the Nation's forest populations. Most inventories at this scale rely on a sampling scheme that has historically been described as a three-step process. In the first step, a set of random points is located in a two-dimensional space, specifically the land area of interest. The second step concerns the selection of a set of observation times while the third step chooses a cluster of trees in the vicinity of each sample point at each observation time. The first step, and only the first step, was viewed as random, leading to a sample design description in which the sample frame partitions the two-dimensional areal population. In Roesch (2008), I addressed the fact that in today's panelized sample designs, the determination of the set of observation times is also random, and the sampled population and the sampling frame are three-dimensional.

To make estimates for the target population, the sampled population must be identifiably associated with the target population. This association requires knowledge of the probability of selection for the realized set of observations on each tree (or element) in the sample over the course of the period of interest. Because there are potentially many sets of observation times realizable for each element in the population, I described the sample unit as a three-dimensional jigsaw puzzle piece resulting from partitioning the three-dimensional population volume. The description meets the requirements for a probability sample: the

[1] Mathematical Statistician, U.S. Forest Service, Southern Research Station, 200 WT Weaver Blvd., Asheville, NC 28804. To contact, call 828-257-4871 or email at froesch@fs.fed.us.

population is divided up into mutually exclusive, exhaustive sample units (the three-dimensional puzzle pieces) that in toto make up the sample frame. Each unit has a definite probability of selection and the total of these probabilities is equal to 1.

AN APPLICATION

Traditionally, many measures associated with forest trees have been reported within tree size classes, such as tree diameter classes. For instance, basal area or volume growth within 2-inch diameter classes for each year within a specific period may be of interest. The contribution of measurement error to total variance is usually large enough to preclude the measurement of the same trees more frequently than about every 5 years. Often the measurement interval and the period of interest are long enough for a large number of trees in the population to grow through multiple diameter classes creating a potentially intractable problem from the viewpoint of successively applied two-dimensional samples. I show below that estimation under the three-dimensional paradigm is both obvious and manageable.

In this application, I use FIA data to estimate annual basal area growth of survivor trees within specific size classes (Table 1) over a defined area (A) and temporal period. FIA conducts a continuous forest inventory using a rotating panel design (Bechtold and Patterson (2005) and Roesch (2007). The design consists of

Table 1.—Diameter classes corresponding to standard merchantability limits with accommodation for differences in hardwood and softwood saw log standards

Diameter Class	Lower Limit (Diameter ≥)	Upper Limit (Diameter <)
D1	0	5
D2	5	7
D3	7	9
D4	9	11
D5	11	∞

g mutually exclusive, spatially disjoint temporal panels. These panels are measured in sequence for g consecutive years, after which the sequence reinitiates. That is, if panel 1 is measured in year y, it will also be measured in years $y + g$, $y + 2g$, and so on. Panel 2 would then be measured in years $y + 1$, $y + 1 + g$, $y + 1 + 2g$, etc. Because FIA adheres to a two-dimensional view of this design, the program groups these data into evaluation groups of g years and then ignores temporal differences in observations within an evaluation group. The interested reader is referred to the “temporally indifferent method” in Patterson and Reams (2005). The temporally indifferent method is a smoothing function that has the tendency to obfuscate temporal trends and delay recognition of those trends. A judicious application of the three-dimensional view of this design can negate the necessity of the temporal indifference assumption and its associated problems.

To fully exploit the three-dimensional view, we must look at the data differently than it has been traditionally viewed. With respect to annual growth, we note that each plot is not only located in a particular place, but that it is also observed at particular times, and that the times of observation are possibly more important than the place of observation, once place is accounted for. Initially, we will focus on two observations for each plot. Assign to each observation of variable x labels for plot i and (adjusted) beginning date t_i^b and ending date t_i^e, separated by the (adjusted) time span of s_i years. Represent each of these observations as x_i^b and x_i^e, respectively. The dates and times are adjusted to approximate the time of observation relative to the proportion of growing season elapsed within a year. Although beyond the scope of this investigation, this could be done using data contributing to the USDA plant hardiness zone maps (USDA 2012). For simplicity, we make two assumptions, both of which can be refined by an appropriate model to suit a particular investigator or alternative application, as needed. The first is that we assume that the growing season spans from March 1 to November 30 everywhere within the area of interest. The second assumption is that growth for each plot is uniform throughout the growing season. We can then represent each observation date as the year of observation plus the proportion of the growing season that has elapsed (i.e., in the format *year.p*), and s_i is simply the difference between the two. Because we have no observations between x_i^b and x_i^e, we make the further assumption that basal area growth for each living tree is uniform between the two observations (e.g., across s_i). This assumption could also be refined by the application of the appropriate model, such as conditioning on x_i^b or on annual precipitation. We then allocate the proportion of basal area growth observed over s_i to the proportion of each year spanned by s_i, (thereby accounting for the marginal probability of the time dimension). This assumption of linear (basal area) growth is an approximation that should only be used for relatively short time intervals. Well-developed growth models would provide better estimates on individual trees, but can be unavailable for many of the species and condition classes encountered in a wide area forest monitoring effort. The assumption that basal area growth is uniform between observations allows us to estimate when the threshold for each diameter class limit was crossed and to allocate growth within diameter classes to the years the growth occurred in those diameter classes. This method is a major advantage over FIA's current estimation methods, because the latter do not provide a mechanism to accomplish this. This development leads immediately to two simple estimators for annual basal area growth (within diameter class), a probability proportional to size estimator (BAG_y^{DCPPS}):

$$BAG_y^{DCPPS} = \frac{1}{n_y}\sum_{i=1}^{n_y}\frac{bag_{i,y}}{p_{i,y}}, \qquad [1]$$

where:
n_y = the number of plots observing growth in year y,
$P_{i,y}$ = the product of portion of year y growing season observed by plot *i* and the portion of plot *i* area within the area of interest, and
$bag_{i,y}$ the basal area growth observed on plot *i* and assignable to year *y*;

and a ratio estimator (BAG_y^{DCRAT}):

$$BAG_y^{DCRAT} = \frac{\sum_{i=1}^{n_y} bag_{i,y}}{\sum_{i=1}^{n_y} p_{i,y}}. \qquad [2]$$

These estimators were used to obtain annual estimates from 2006 to 2010 of basal area growth of survivor trees per acre from FIA data for South Carolina. The results are compared to FIA's end of period estimator (EOP) and an improved diameter class estimator (DC) (Sheffield and Turner, 2010), both of which are based on the temporal indifference assumption.

RESULTS

Figure 1 gives the results for estimating the annual basal area growth of survivor trees by the diameter classes given in Table 1 from each of the four estimators. The figure shows the results for the PPS estimator (eq. 1) (top left), the ratio estimator (eq. 2) (top right), the pooled EOP estimator (bottom left), and the pooled diameter class (DC) estimator (bottom right). Note that under a non-stringent condition, the estimators resulting in the top two graphs are unbiased. A linear trend for the intervals covering the year of interest is sufficient for unbiasedness. Note also that these two estimators gave almost the same results, which are quite different from the estimators resulting in the bottom two graphs. For the EOP estimators to be unbiased, a flat line trend (i.e., linear with a slope of 0) over all years used in the estimators would have to exist. With a 5-year cycle, a flat line trend must have been true for the 10 years before any annual EOP estimate of growth. From the top graphs, we see that a flat line trend is definitely not indicated for three of the five diameter classes. Between 2006 and 2010, the pooled DC (Sheffield) estimator gave results that were at times closer to the results for the estimators in the top two graphs than the FIA EOP estimator, in the bottom left graph, although the trend through those years is not discernible in either of the EOP estimators.

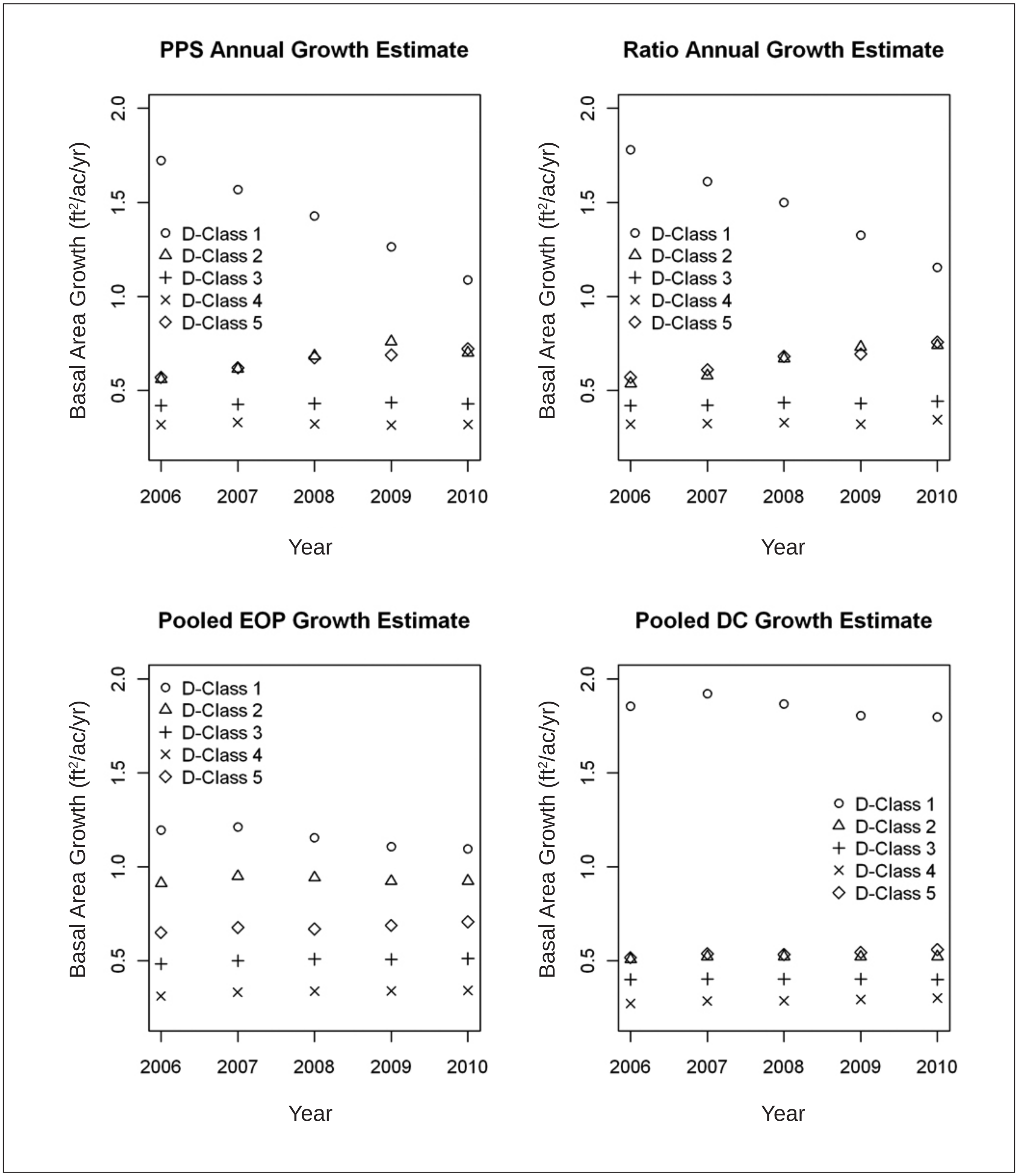

Figure 1.—Annual basal area growth estimates by the diameter classes given in Table 1 from each of the four estimators, clockwise from the top left: the probability proportional to size (PPS) estimator, the ratio estimator, the pooled diameter class (Sheffield) estimator, and the pooled EOP (FIA-temporally indifferent) estimator.

CONCLUSIONS

The field of statistics gives us many estimation tools to bolster analyses. All four estimators discussed here somehow use "outside information" to make annual estimates. In the estimators resulting in the top two graphs of Figure 1, the "outside information" has a clear relationship to the estimates of interest, because the observations span the estimates to which the data contribute. It is clear in the formulation of the EOP estimators and in Figure 2, that much of the outside information used in those estimators does not span the time estimated. In the results for 2010, for instance, Figure 2 shows that 80 percent of the information used in the EOP estimators is from "outside" of, and prior to, 2010. This prior information is incorporated under a model that assumes that the mean of the "outside" information is the same as the mean for 2010. It is not very helpful to start a search for trend by first assuming that there is not any trend. The description of continuous forest inventories as a sample of a three-dimensional population is uniquely informative. It arose from the recognition of the importance of the time of observation on the outcome of the sample and it is useful for putting temporally ordered observations into perspective while formulating intuitively appealing model-unbiased estimators of growth and trend.

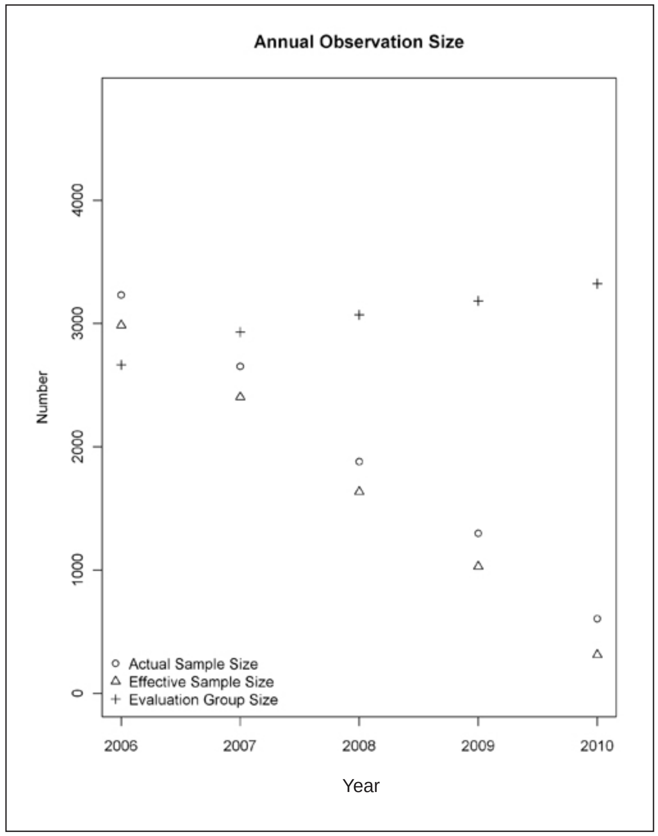

Figure 2.—The observation sizes of annual growth estimates for each year. The actual sample size is the number of plots with growth intervals spanning at least part of the year. The effective sample size is the sum of plot-year proportions covered by the plots making up the actual sample size. The evaluation group size is the number of plots used in the FIA EOP estimator (temporally indifferent) and the Sheffield DC EOP estimator. The majority of the plots used in the two pooled estimators (FIA and Sheffield) do not observe any part of the growth interval for that year.

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.

Patterson, P.L.; Reams, G.A. 2005. **Combining panels for Forest Inventory and Analysis estimation.** In: Bechtold, W.A.; Patterson, P.L., eds. 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: 11-26.

Roesch, F.A. 2007. **Compatible estimators of the components of change for a rotating panel forest inventory design.** Forest Science. 53(1): 50-61.

Roesch, F.A. 2008. **An alternative view of continuous forest inventories.** Forest Science. 54(4): 455-464.

Roesch, F.A., Jr.; Green, E.J.; Scott, C.T. 1993. **An alternative view of forest sampling.** Survey Methodology. 19(2): 199-204.

Sheffield, R.M.; Turner, J.A. 2010. **Improving FIA diameter class growth distributions.** Book of Abstracts. 2010 Forest Inventory and Analysis symposium, Monitoring Across Borders. 2010 October 5-7; Knoxville, TN.

USDA. 2012. **USDA Plant Hardiness Zone Map, 2012.** Agricultural Research Service, U.S. Department of Agriculture. Accessed from http://planthardiness.ars.usda.gov. [Date accessed unknown].

The content of this paper reflects the views of the author(s), who are responsible for the facts and accuracy of the information presented herein.